



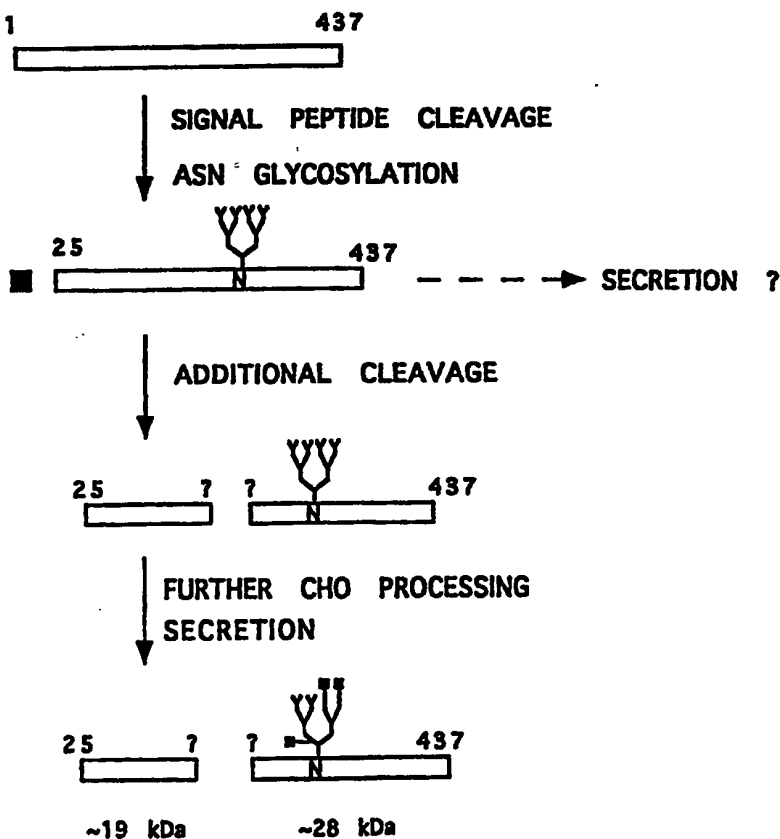
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, C07K 14/46, 14/435, A61K 48/00		A1	(11) International Publication Number: WO 95/18856
			(43) International Publication Date: 13 July 1995 (13.07.95)
(21) International Application Number: PCT/US94/14992		(81) Designated States: AU, CA, JP, KR, NZ, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) International Filing Date: 30 December 1994 (30.12.94)		Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(30) Priority Data:			
08/176,427	30 December 1993 (30.12.93) US		
08/356,060	14 December 1994 (14.12.94) US		
(71) Applicants: PRESIDENT AND FELLOWS OF HARVARD COLLEGE [US/US]; 124 Mt. Auburn Street, Cambridge, MA 02138 (US). IMPERIAL CANCER RESEARCH TECHNOLOGY, LTD. [GB/GB]; Sardinia House, Sardinia Street, London WC2A 3NL (GB).			
(72) Inventors: INGHAM, Philip, W.; 83 Middle Way Road, Summertown, Oxford OX2 7LE (GB). MCMAHON, Andrew, P.; 128 Kendall Road, Lexington, MA 02173 (US). TABIN, Clifford, J.; 102 Hancock Street, Cambridge, MA 02138 (US).			
(74) Agents: VINCENT, Matthew, P. et al.; Lahive & Cockfield, 60 State Street, Boston, MA 02109 (US).			

(54) Title: VERTEBRATE EMBRYONIC PATTERN-INDUCING HEDGEHOG-LIKE PROTEINS

(57) Abstract

The present invention concerns the discovery that proteins encoded by a family of vertebrate genes, termed here *hedgehog*-related genes, comprise morphogenic signals produced by embryonic patterning centers, and are involved in the formation of ordered spatial arrangements of differentiated tissues in vertebrates. The present invention makes available compositions and methods that can be utilized, for example to generate and/or maintain an array of different vertebrate tissue both *in vitro* and *in vivo*.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LJ	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

VERTEBRATE EMBRYONIC PATTERN-INDUCING HEDGEHOG-LIKE PROTEINS.

Background of the Invention

5 Pattern formation is the activity by which embryonic cells form ordered spatial arrangements of differentiated tissues. The physical complexity of higher organisms arises during embryogenesis through the interplay of cell-intrinsic lineage and cell-extrinsic signaling. Inductive interactions are essential to embryonic patterning in vertebrate development from the earliest establishment of the body plan, to the patterning of the organ
10 systems, to the generation of diverse cell types during tissue differentiation (Davidson, E., (1990) *Development* 108: 365-389; Gurdon, J. B., (1992) *Cell* 68: 185-199; Jessell, T. M. et al., (1992) *Cell* 68: 257-270). The effects of developmental cell interactions are varied. Typically, responding cells are diverted from one route of cell differentiation to another by inducing cells that differ from both the uninduced and induced states of the responding cells
15 (inductions). Sometimes cells induce their neighbors to differentiate like themselves (homoio-genetic induction); in other cases a cell inhibits its neighbors from differentiating like itself. Cell interactions in early development may be sequential, such that an initial induction between two cell types leads to a progressive amplification of diversity. Moreover, inductive interactions occur not only in embryos, but in adult cells as well, and can act to
20 establish and maintain morphogenetic patterns as well as induce differentiation (J.B. Gurdon (1992) *Cell* 68:185-199).

 The origin of the nervous system in all vertebrates can be traced to the end of gastrulation. At this time, the ectoderm in the dorsal side of the embryo changes its fate from epidermal to neural. The newly formed neuroectoderm thickens to form a flattened structure
25 called the neural plate which is characterized, in some vertebrates, by a central groove (neural groove) and thickened lateral edges (neural folds). At its early stages of differentiation, the neural plate already exhibits signs of regional differentiation along its anterior posterior (A-P) and mediolateral axis (M-L). The neural folds eventually fuse at the dorsal midline to form the neural tube which will differentiate into brain at its anterior end and spinal cord at its
30 posterior end. Closure of the neural tube creates dorsal/ventral differences by virtue of previous mediolateral differentiation. Thus, at the end of neurulation, the neural tube has a clear anterior-posterior (A-P), dorsal ventral (D-V) and mediolateral (M-L) polarities (see, for example, *Principles in Neural Science* (3rd), eds. Kandel, Schwartz and Jessell, Elsevier Science Publishing Company: NY, 1991; and *Developmental Biology* (3rd), ed. S.F. Gilbert,
35 Sinauer Associates: Sunderland MA, 1991). Inductive interactions that define the fate of cells within the neural tube establish the initial pattern of the embryonic vertebrate nervous system. In the spinal cord, the identify of cell types is controlled, in part, by signals from two

midline cell groups, the notochord and floor plate, that induce neural plate cells to differentiate into floor plate, motor neurons, and other ventral neuronal types (van Straaten et al. (1988) *Anat. Embryol.* 177:317-324; Placzek et al. (1993) *Development* 117:205-218; Yamada et al. (1991) *Cell* 64:635-647; and Hatta et al. (1991) *Nature* 350:339-341). In addition, signals from the floor plate are responsible for the orientation and direction of commissural neuron outgrowth (Placzek, M. et al., (1990) *Development* 110: 19-30). Besides patterning the neural tube, the notochord and floorplate are also responsible for producing signals which control the patterning of the somites by inhibiting differentiation of dorsal somite derivatives in the ventral regions (Brand-Saberi, B. et al., (1993) *Anat. Embryol.* 188: 239-245; Porquie, O. et al., (1993) *Proc. Natl. Acad. Sci. USA* 90: 5242-5246).

Another important signaling center exists in the posterior mesenchyme of developing limb buds, called the Zone of Polarizing Activity, or "ZPA". When tissue from the posterior region of the limb bud is grafted to the anterior border of a second limb bud, the resultant limb will develop with additional digits in a mirror-image sequence along the anteroposterior axis (Saunders and Gasseling, (1968) *Epithelial-Mesenchymal Interaction*, pp. 78-97). This finding has led to the model that the ZPA is responsible for normal anteroposterior patterning in the limb. The ZPA has been hypothesized to function by releasing a signal, termed a "morphogen", which forms a gradient across the early embryonic bud. According to this model, the fate of cells at different distances from the ZPA is determined by the local concentration of the morphogen, with specific thresholds of the morphogen inducing successive structures (Wolpert, (1969) *Theor. Biol.* 25:1-47). This is supported by the finding that the extent of digit duplication is proportional to the number of implanted ZPA cells (Tickle, (1981) *Nature* 254:199-202).

A candidate for the putative ZPA morphogen was identified by the discovery that a source of retinoic acid can result in the same type of mirror-image digit duplications when placed in the anterior of a limb bud (Tickle et al., (1982) *Nature* 296:564-565; Summerbell, (1983) *J. Embryol.* 78:269-289). The response to exogenous retinoic acid is concentration dependent as the morphogen model demands (Tickle et al., (1985) *Dev. Biol.* 109:82-95). Moreover, a differential distribution of retinoic acid exists across the limb bud, with a higher concentration in the ZPA region (Thaller and Eichele, (1987) *Nature* 327:625-628).

Recent evidence, however, has indicated that retinoic acid is unlikely to be the endogenous factor responsible for ZPA activity (reviewed in Brockes, (1991) *Nature* 350:15; Tabin, (1991) *Cell* 66:199-217). It is now believed that rather than directly mimicking an endogenous signal, retinoic acid implants act by inducing an ectopic ZPA. The anterior limb tissue just distal to a retinoic acid implant and directly under the ectoderm has been demonstrated to acquire ZPA activity by serially transplanting that tissue to another limb bud (Summerbell and Harvey, (1983) *Limb Development and Regeneration* pp. 109-118; Wanek

et al., (1991) *Nature* 350:81-83). Conversely, the tissue next to a ZPA graft does not gain ZPA activity (Smith, (1979) *J. Embryol* 52:105-113). Exogenous retinoic acid would thus appear to act upstream of the ZPA in limb patterning.

The immediate downstream targets of ZPA action are not known. However, one important set of genes which are ectopically activated during ZPA-induced pattern duplications are the 5' genes of the Hoxd cluster. These genes are normally expressed in a nested pattern emanating from the posterior margin of the limb bud (Dolle et al., (1989) *Nature* 342:767-772; Izpisua-Belmonte et al., (1991) *Nature* 350:585-589). This nested pattern of Hox gene expression has been directly demonstrated to determine the identity of the structures produced along the anteroposterior axis of the limb (Morgan et al., (1993) *Nature* 358:236-239). As this would predict, ZPA grafts which produce mirror-image duplication of structures at an anatomical level first lead to the ectopic activation of the Hoxd genes in a mirror-image duplication at the molecular level. (Nohno et al., (1991) *Cell* 64:1197-1205; Izpisua-Belmonte et al., (1991) *Nature* 350:585-589). The molecular signals which regulate the expression of these important genes are currently not understood.

Summary of the Invention

The present invention relates to the discovery of a novel family of proteins present in vertebrate organisms, referred to hereinafter as "*hedgehog*" proteins, which proteins have apparent broad involvement in the formation and maintenance of ordered spatial arrangements of differentiated tissues in vertebrates, and can be used to generate and/or maintain an array of different vertebrate tissue both *in vitro* and *in vivo*.

In general, the invention features *hedgehog* polypeptides, preferably substantially pure preparations of one or more of the subject *hedgehog* polypeptides. The invention also provides recombinantly produced *hedgehog* polypeptides. In preferred embodiments the polypeptide has a biological activity including: an ability to modulate proliferation, survival and/or differentiation of mesodermally-derived tissue, such as tissue derived from dorsal mesoderm; the ability to modulate proliferation, survival and/or differentiation of ectodermally-derived tissue, such as tissue derived from the neural tube, neural crest, or head mesenchyme; the ability to modulate proliferation, survival and/or differentiation of endodermally-derived tissue, such as tissue derived from the primitive gut. Moreover, in preferred embodiments, the subject *hedgehog* proteins have the ability to induce expression of secondary signaling molecules, such as members of the Transforming Growth Factor β family, as well as members of the fibroblast growth factor (FGF) family.

In a certain embodiments, the polypeptide is identical with or homologous to a *Sonic hedgehog* (*Shh*) polypeptide, such as a mammalian *Shh* represented by SEQ ID Nos:13 or 11, an avian *Shh* represented by SEQ ID No: 8, or a fish *Shh* represented by SEQ ID No: 12. For instance, the *Shh* polypeptide preferably has an amino acid sequence at least 70% homologous to a polypeptide represented by any of SEQ ID Nos: 8, 11, 12 or 13, though polypeptides with higher sequence homologies of, for example, 80%, 90% or 95% are also contemplated. Exemplary *Shh* proteins are represented by SEQ ID No. 40. The *Shh* polypeptide can comprise a full length protein, such as represented in the sequence listings, or it can comprise a fragment of, for instance, at least 5, 10, 20, 50, 100 or 150 amino acids in length. Preferred *hedgehog* polypeptides include *Shh* sequences corresponding approximately to the natural proteolytic fragments of the *hedgehog* proteins, such as from about Cys-24 through Glu-188, or from about Asn-189 through Ala-475 of the human *Shh* protein, or analogous fragments thereto.

In other embodiments, the polypeptide is identical with or homologous to an *Indian hedgehog* (*Ihh*) polypeptide, such as a human *Ihh* represented by SEQ ID No:14, or a mouse *Ihh* represented by SEQ ID No: 10. For instance, the *Ihh* polypeptide preferably has an amino acid sequence at least 70% homologous to a polypeptide represented by either of SEQ ID Nos: 10 or 14, though *Ihh* polypeptides with higher sequence homologies of, for example, 80%, 90% or 95% are also contemplated. The polypeptide can comprise the full length protein represented by in part by these sequences, or it can comprise a fragment of, for instance, at least 5, 10, 20, 50, 100 or 150 amino acids in length. Preferred *Ihh* polypeptides comprise an N-terminal fragment including Arg-1 through Glu-94, or a C-terminal fragment including His-95 through Ser-3312 of the human *Ihh* represented by SEQ ID No: 14, or analogous fragments thereto.

In still further embodiments, the polypeptide is identical with or homologous to a *Desert hedgehog* (*Dhh*) polypeptide, such as a mouse *Dhh* represented by SEQ ID No: 9. For instance, the *Dhh* polypeptide preferably has an amino acid sequence at least 70% homologous to a polypeptide represented by SEQ ID No: 9, though *Dhh* polypeptides with higher sequence homologies of, for example, 80%, 90% or 95% are also contemplated. The polypeptide can comprise the full length protein represented by this sequence, or it can comprise a fragment of, for instance, at least 5, 10, 20, 50, 100 or 150 amino acids in length. Preferred *Dhh* polypeptides comprise *Dhh* sequences corresponding to the N-terminal portion of the protein, e.g. Cys-23 through Asp-189 or Asn-190 through Gly-396 of SEQ ID No: 9, or analogous fragments thereto.

Moreover, as described below, the *hedgehog* polypeptide can be either an agonist (e.g. mimics), or alternatively, an antagonist of a biological activity of a naturally occurring form of the protein, e.g., the polypeptide is able to modulate differentiation and/or growth

5

and/or survival of a cell responsive to authentic *hedgehog* proteins. Homologs of the subject *hedgehog* proteins include versions of the protein which are resistant to proteolytic cleavage, as for example, due to mutations which alter potential cleavage sequences or which inactivate an enzymatic activity associated with the protein.

5 The *hedgehog* polypeptides of the present invention can be glycosylated, or conversely, by choice of the expression system or by modification of the protein sequence to preclude glycosylation, reduced carbohydrate analogs can also be provided. Glycosylated forms include derivatization with glycosaminoglycan chains. Likewise, *hedgehog* polypeptides can be generated which lack an endogenous signal sequence (though this is
10 typically cleaved off even if present in the pro-form of the protein).

 The subject proteins can also be provided as chimeric molecules, such as in the form of fusion proteins. For instance, the *hedgehog* protein can be provided as a recombinant fusion protein which includes a second polypeptide portion, e.g., a second polypeptide having an amino acid sequence unrelated to *hedgehog*, e.g. the second polypeptide portion is
15 glutathione-S-transferase, e.g. the second polypeptide portion is an enzymatic activity such as alkaline phosphatase, e.g. the second polypeptide portion is an epitope tag.

 Yet another aspect of the present invention concerns an immunogen comprising a *hedgehog* polypeptide in an immunogenic preparation, the immunogen being capable of eliciting an immune response specific for a *hedgehog* polypeptide; e.g. a humoral response,
20 e.g. an antibody response; e.g. a cellular response. In preferred embodiments, the immunogen comprising an antigenic determinant, e.g. a unique determinant, from a protein represented by one of SEQ ID Nos. 8-14.

 A still further aspect of the present invention features antibodies and antibody preparations specifically reactive with an epitope of the *hedgehog* immunogen.

25 Another aspect of the present invention provides a substantially isolated nucleic acid having a nucleotide sequence which encodes a *hedgehog* polypeptide. In preferred embodiments, the encoded polypeptide specifically agonizes or antagonizes inductive events mediated by wild-type *hedgehog* proteins. The coding sequence of the nucleic acid can comprise a sequence which is identical to a coding sequence represented in one of SEQ ID
30 Nos: 1-7, or it can merely be homologous to one or more of those sequences. For instance, the *hedgehog* encoding sequence preferably has a sequence at least 70% homologous to a nucleotide sequence in one or more of SEQ ID Nos: 1-7, though higher sequence homologies of, for example, 80%, 90% or 95% are also contemplated. The polypeptide encoded by the nucleic acid can comprise an amino acid sequence represented in one of SEQ ID Nos: 8-14
35 such as one of those full length proteins, or it can comprise a fragment of that nucleic acid, which fragment may, for instance, encode a fragment which is, for example, at least 5, 10, 20,

6

50 or 100 amino acids in length. The polypeptide encoded by the nucleic acid can be either an agonist (e.g. mimics), or alternatively, an antagonist of a biological activity of a naturally occurring form of a *hedgehog* protein.

Furthermore, in certain preferred embodiments, the subject *hedgehog* nucleic acid will include a transcriptional regulatory sequence, e.g. at least one of a transcriptional promoter or transcriptional enhancer sequence, which regulatory sequence is operably linked to the *hedgehog* gene sequence. Such regulatory sequences can be used in to render the *hedgehog* gene sequence suitable for use as an expression vector.

In yet a further preferred embodiment, the nucleic acid hybridizes under stringent conditions to a nucleic acid probe corresponding to at least 12 consecutive nucleotides of one or more of SEQ ID Nos: 1-7; though preferably corresponding to at least 20 consecutive nucleotides; and more preferably corresponding to at least 40, 50 or 75 consecutive nucleotides of one or more of SEQ ID Nos: 1-7.

The invention also features transgenic non-human animals, e.g. mice, rats, rabbits, chickens, frogs or pigs, having a transgene, e.g., animals which include (and preferably express) a heterologous form of a *hedgehog* gene described herein, or which misexpress an endogenous *hedgehog* gene, e.g., an animal in which expression of one or more of the subject *hedgehog* proteins is disrupted. Such a transgenic animal can serve as an animal model for studying cellular and tissue disorders comprising mutated or mis-expressed *hedgehog* alleles or for use in drug screening.

The invention also provides a probe/primer comprising a substantially purified oligonucleotide, wherein the oligonucleotide comprises a region of nucleotide sequence which hybridizes under stringent conditions to at least 10 consecutive nucleotides of sense or antisense sequence of SEQ ID No: 1, or naturally occurring mutants thereof. Nucleic acid probes which are specific for each of the classes of vertebrate *hedgehog* proteins are contemplated by the present invention, e.g. probes which can discern between nucleic acid encoding an *Shh* versus an *Ihh* versus a *Dhh* versus an *Mhh*. In preferred embodiments, the probe/primer further includes a label group attached thereto and able to be detected. The label group can be selected, e.g., from a group consisting of radioisotopes, fluorescent compounds, enzymes, and enzyme co-factors. Probes of the invention can be used as a part of a diagnostic test kit for identifying dysfunctions associated with mis-expression of a *hedgehog* protein, such as for detecting in a sample of cells isolated from a patient, a level of a nucleic acid encoding a subject *hedgehog* protein; e.g. measuring a *hedgehog* mRNA level in a cell, or determining whether a genomic *hedgehog* gene has been mutated or deleted. Preferably, the oligonucleotide is at least 10 nucleotides in length, though primers of 20, 30, 50, 100, or 150 nucleotides in length are also contemplated.

In yet another aspect, the invention provides an assay for screening test compounds for inhibitors, or alternatively, potentiators, of an interaction between a *hedgehog* protein and a *hedgehog* receptor. An exemplary method includes the steps of (i) combining a *hedgehog* receptor, either soluble or membrane bound (including whole cells), a *hedgehog* polypeptide, and a test compound, e.g., under conditions wherein, but for the test compound, the *hedgehog* protein and the *hedgehog* receptor are able to interact; and (ii) detecting the formation of a complex which includes the *hedgehog* protein and the receptor either by directly quantitating the complex or by measuring inductive effects of the *hedgehog* protein. A statistically significant change, such as a decrease, in the formation of the complex in the presence of a test compound (relative to what is seen in the absence of the test compound) is indicative of a modulation, e.g., inhibition, of the interaction between the *hedgehog* protein and the receptor.

Another aspect of the present invention relates to a method of inducing and/or maintaining a differentiated state, causing proliferation, and/or enhancing survival of a cell (from a vertebrate organism) responsive to a *hedgehog* protein, by contacting the cells with a *hedgehog* agonist. For example, the present method is applicable to cell culture technique, such as in the culturing of neuronal and other cells whose survival or differentiative state is dependent on *hedgehog* function. Moreover, *hedgehog* agonists and antagonists can be used for therapeutic intervention, such as to enhance survival and maintenance of neurons and other neural cells in both the central nervous system and the peripheral nervous system, as well as to influence other vertebrate organogenic pathways, such as other ectodermal patterning, as well as certain mesodermal and endodermal differentiation processes. In addition to the vertebrate *hedgehog*-like proteins, the present invention further contemplates the use of *Drosophila Hedgehog* (Dros-HH) to induce cells and tissue of vertebrate organisms in similar fashion to the subject *hedgehog* proteins.

Another aspect of the present invention provides a method of determining if a subject, e.g. a human patient, is at risk for a disorder characterized by unwanted cell proliferation or aberrant control of differentiation. The method includes detecting, in a tissue of the subject, the presence or absence of a genetic lesion characterized by at least one of (i) a mutation of a gene encoding a *hedgehog* protein, e.g. represented in SEQ ID No: 2, or a homolog thereof; or (ii) the mis-expression of a *hedgehog* gene. In preferred embodiments, detecting the genetic lesion includes ascertaining the existence of at least one of: a deletion of one or more nucleotides from a *hedgehog* gene; an addition of one or more nucleotides to the gene, a substitution of one or more nucleotides of the gene, a gross chromosomal rearrangement of the gene; an alteration in the level of a messenger RNA transcript of the gene; the presence of a non-wild type splicing pattern of a messenger RNA transcript of the gene; or a non-wild type level of the protein.

8

For example, detecting the genetic lesion can include (i) providing a probe/primer including an oligonucleotide containing a region of nucleotide sequence which hybridizes to a sense or antisense sequence of a *hedgehog* gene, e.g. a nucleic acid represented in one of SEQ ID Nos: 1-7, or naturally occurring mutants thereof, or 5' or 3' flanking sequences naturally associated with the *hedgehog* gene; (ii) exposing the probe/primer to nucleic acid of the tissue; and (iii) detecting, by hybridization of the probe/primer to the nucleic acid, the presence or absence of the genetic lesion; e.g. wherein detecting the lesion comprises utilizing the probe/primer to determine the nucleotide sequence of the *hedgehog* gene and, optionally, of the flanking nucleic acid sequences. For instance, the probe/primer can be employed in a polymerase chain reaction (PCR) or in a ligation chain reaction (LCR). In alternate embodiments, the level of a *hedgehog* protein is detected in an immunoassay using an antibody which is specifically immunoreactive with the *hedgehog* protein.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See, for example, *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); *DNA Cloning*, Volumes I and II (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis et al. U.S. Patent No: 4,683,195; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Transcription And Translation* (B. D. Hames & S. J. Higgins eds. 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); *Methods In Enzymology*, Vols. 154 and 155 (Wu et al. eds.), *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds., 1986); *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986).

Other features and advantages of the invention will be apparent from the following detailed description, and from the claims.

Brief Description of the Drawings

Figure 1 represents the amino acid sequences of two chick *hh* clones, chicken *hedgehog*-A (pCHA; SEQ ID No:35) and chicken *hedgehog*-B (pCHB; SEQ ID No:36). These clones were obtained using degenerate primers corresponding to the underlined amino

acid residues of the *Drosophila* sequence (corresponding to residues 161-232 of SEQ ID No:34) also shown in Figure 1, followed by nested PCR using chicken genomic DNA.

Figure 2 is an alignment comparing the amino acid sequences of chick *Shh* (SEQ ID No:8) with its *Drosophila* homolog (SEQ ID No:34). *Shh* residues 1-26 correspond to the proposed signal peptide. Identical residues are enclosed by boxes and gaps in order to highlight similarity. The nucleotide sequence of *Shh* has been submitted to Genbank.

Figure 3 is a hydropathy plot for the predicted chick *Shh* protein, generated by the methods of Kyte and Doolittle (1982). The values of hydrophobicity are plotted against the amino acid positions. Negative values predict a hydrophobic domain of the protein.

Figure 4 is an alignment comparing the amino acid sequences of various *hh* proteins. The white region on the amino terminus of chicken *Shh* corresponds to the putative signal peptide. The black box refers to a highly conserved region from aa residues 26-207 of SEQ ID No:8). The arrows point to exon boundaries in the *Drosophila* gene (Lee et al. (1992) *Cell* 71: 33-50). In each case, the proteins are compared to chicken *Shh* (SEQ ID No:8) and the percent amino acid identity is indicated in each region's box.

Figure 5A is a "pileup" alignment of predicted amino acid sequences which compares *Drosophila hh* (D-*hh*; SEQ ID No:34), mouse *hh* (M-*Dhh*; SEQ ID No:9; M-*Ihh*; SEQ ID No:10; M-*Shh*; SEQ ID No:11), chicken *hh* (C-*Shh*; SEQ ID No:8), and zebrafish *hh* (Z-*Shh*; SEQ ID No:12). The predicted hydrophobic transmembrane/signal sequences are indicated in italics and the predicted signal sequence processing site is arrowed. The positions of introns interrupting the *Drosophila hh* and M-*Dhh* open reading frames are indicated by arrowheads. All amino acids shared among the six predicted *hh* proteins are indicated in bold. Figure 5B is a sequence alignment of the N-terminal portion of vertebrate *hedgehog* proteins, and the predicted degenerate sequence "CON" (SEQ ID No: 41).

Figure 6 is an inter- and cross-species comparison of amino acid identities among the predicted processed *hh* proteins shown in Figure 5A. All values are percentages. Figures in parentheses represent similarities allowing for conservative amino acid substitutions.

Figure 7 is a representation of the DNA constructs used in transgenic studies to study ectopic expression of chick *Shh* in mouse embryos. Constructs were generated for ectopic expression of cDNA clones in the *Wnt-1* expression domain and tested in transgenic mice embryos using a lac-Z reporter (pWEXP-lacZ (used as a control)) and a chick *Shh* reporter (pWEXP-CShh). The pWEXP-CShh construct contained two tandem head to tail copies of a chick *Shh* cDNA. The results of WEXP2-CShh transgenic studies are shown in Table 1.

Figure 8 is a model for anterioposterior limb patterning and the Zone of Polarizing Activity (ZPA), based on Saunders and Gasseling (1968). The left portion of the diagram

10

schematizes a stage 20 limb bud. The somites are illustrated as blocks along the left margin of the limb bud; right portion of the same panel illustrates the mature wing. The hatched region on the posterior limb is the ZPA. Normally, the developed wing contains three digits II, III, and IV. The figure further shows the result of transplanting a ZPA from one limb bud to the anterior margin of another. The mature limb now contains six digits IV, III, II, II, III, and IV in a mirror-image duplication of the normal pattern. The large arrows in both panels represent the signal produced by the ZPA which acts to specify digit identity.

Figures 9A and 9B illustrate the comparison of zebrafish *Shh* (Z-*Shh*) and *Drosophila hh* (hh) amino acid sequences. Figure 9A is an alignment of zebrafish *Shh* and *Drosophila hh* amino acid sequences. Identical amino acids are linked by vertical bars. Dots indicate gaps introduced for optimal alignment. Putative transmembrane/signal peptide sequences are underlined (Kyte and Doolittle (1982) *J Mol Biol* 157:133-148). The position of exon boundaries in the *Drosophila* gene are indicated by arrowheads. The region of highest similarity between Z-*Shh* and *hh* overlaps exon 2. Figure 9B is a schematic comparison of Z-*Shh* and *Drosophila hh*. Black boxes indicate the position of the putative transmembrane/signal peptide sequences, relative to the amino-terminus. Sequence homologies were scored by taking into account the alignment of chemically similar amino acids and percentage of homology in the boxed regions is indicated.

Figure 10 is an alignment of partial predicted amino acid sequences from three different zebrafish *hh* homologs. One of these sequences corresponds to *Shh*, while the other two define additional *hh* homologs in zebrafish, named *hh(a)* and *hh(b)*. Amino acid identities among the three partial homologs are indicated by vertical bars.

Figure 11 is a schematic representations of chick and mouse *Shh* proteins. The putative signal peptides and Asn-linked glycosylation sites are shown. The numbers refer to amino acid positions.

Figure 12 is a schematic representation of myc-tagged *Shh* constructs. The positions of the c-myc epitope tags are shown, as is the predicted position of the proteolytic cleavage site. The shaded area following the signal peptide of the carboxy terminal tagged construct represents the region included in the Glutathione-S-transferase fusion protein used to generate antisera in rabbits.

Figure 13 is a schematic diagram of *Shh* processing. Illustrated are cleavage of the signal peptide (black box), glycosylation at the predicted Asn residue (N), and the secondary proteolytic cleavage. The question marks indicate that the precise site of proteolytic cleavage has not been determined. The different symbols representing the carbohydrate moiety indicated maturation of this structure in the Golgi apparatus. The dashed arrow leading from

11

the signal peptide cleaved protein indicates that secretion of this species may be an artifact of the incomplete proteolytic processing of *Shh* seen in *Xenopus* oocytes and cos cells.

Figure 14 is a schematic diagram of a model for the coordinated growth and patterning of the limb. *Sonic* is proposed to signal directly to the mesoderm to induce expression of the *Hoxd* and *Bmp-2* genes. The induction of these mesodermal genes requires competence signals from the overlying AER. One such signal is apparently *Fgf-4*. Expression of *Fgf-4* in the AER can be induced by *Sonic* providing an indirect signaling pathway from *Sonic* to the mesoderm. FGFs also maintain expression of *Sonic* in the ZPA, thereby completing a positive feedback loop which controls the relative positions of the signaling centers. While *Fgf-4* provides competence signals to the mesoderm, it also promotes mesodermal proliferation. Thus patterning of the mesoderm is dependent on the same signals which promote its proliferation. This mechanism inextricably integrates limb patterning with outgrowth.

Figure 15 is a schematic diagram of patterning of the *Drosophila* and vertebrate gut. Regulatory interactions responsible for patterning of *Drosophila* midgut (A) are compared to a model for patterning of the vertebrate hindgut (B) based on expression data. Morphologic regional distinctions are indicated to the left (A and B), genes expressed in the visceral mesoderm are in the center panel, those in the gut luminal endoderm are on the right. *HOM/Hox* gene expression domains are boxed. Regionally expressing secreted gene products are indicated by lines. Arrows indicate activating interactions, barred lines, inhibiting interactions. Regulatory interactions in *Drosophila* gut (A) have been established by genetic studies except for the relationship between *dpp* and *hedgehog*, which is hypothesized based on their interactions in the *Drosophila* imaginal discs, *hedgehog* appears to be a signal from the endoderm to the mesoderm, and that *dpp* is expressed in the mesoderm.

Figure 16 is a schematic diagram of chromosomal locations of *Ihh*, *Shh* and *Dhh* in the mouse genome. The loci were mapped by interspecific backcross analysis. The segregation patterns of the loci and flanking genes in backcross animals that were typed for all loci are shown above the chromosome maps. For individual pairs of loci more animals were typed. Each column represents the chromosome identified in the backcross progeny that was inherited from the (C57BL/6J x *M. spretus*) F1 parent. The shaded boxes represent the presence of a C57BL/6J allele and white boxes represent the presence of a *M. spretus* allele. The number of the offsprings inheriting each type of chromosome is listed at the bottom of each column. Partial chromosome linkage maps showing location of *Ihh*, *Shh* and *Dhh* in relation too linked genes is shown. The number of recombinant N₂ animals is

presented over total number of N₂ animals typed to the left of the chromosome maps between each pair of loci. The recombinant frequencies, expressed as genetic distance in centimorgans (\pm one standard error) are also shown. When no recombination between loci was detected, the upper 95% confidence limit of the recombination distance is indicated in parentheses. Gene order was determined by minimizing the number of recombinant events required to explain the allele distribution patterns. The position of loci in human chromosomes can be obtained from GDB (Genome Data Base), a computerized database of human linkage information maintained by the William H. Welch Medical Library of the John Hopkins University (Baltimore, MD).

10

Detailed Description of the Invention

Embryonic inductive signals are key regulatory proteins that function in vertebrate pattern formation, and are present in important signaling centers known to operate embryonically to define the organization of the vertebrate embryo. For example, these signaling structures include the notochord, a transient structure which initiates the formation of the nervous system and helps to define the different types of neurons within it. The notochord also regulates mesodermal patterning along the body axis. Another distinct group of cells having apparent signaling activity is the floorplate of the neural tube (the precursor of the spinal cord and brain) which also signals the differentiation of different nerve cell types. It is also generally believed that the region of mesoderm at the bottom of the buds which form the limbs (called the Zone of Polarizing Activity or ZPA) operates as a signaling center by secreting a morphogen which ultimately produces the correct patterning of the developing limbs.

The present invention concerns the discovery that proteins encoded by a family of vertebrate genes, termed here *hedgehog*-related genes, comprise the signals produced by these embryonic patterning centers. As described herein, each of the disclosed vertebrate *hedgehog* (*hh*) homologs exhibits spatially and temporally restricted expression domains indicative of important roles in embryonic patterning. For instance, the results provided below indicate that vertebrate *hh* genes are expressed in the posterior limb bud, Hensen's node, the early notochord, the floor plate of the neural tube, the fore- and hindgut and their derivatives. These are all important signaling centers known to be required for proper patterning of surrounding embryonic tissues.

The *Hedgehog* family of vertebrate inter-cellular signaling molecules provided by the present invention consists of at least four members. Three of these members, herein referred to as Desert *hedgehog* (*Dhh*), Sonic *hedgehog* (*Shh*) and Indian *hedgehog* (*Ihh*), exist in all vertebrates, including fish, birds, and mammals. A fourth member, herein referred to as

1,3

Moonrat *hedgehog* (*Mhh*), appears specific to fish. According to the appended sequence listing, (see also Table 1) a chicken *Shh* polypeptide is encoded by SEQ ID No:1; a mouse *Dhh* polypeptide is encoded by SEQ ID No:2; a mouse *Ihh* polypeptide is encoded by SEQ ID No:3; a mouse *Shh* polypeptide is encoded by SEQ ID No:4 a zebrafish *Shh* polypeptide is encoded by SEQ ID No:5; a human *Shh* polypeptide is encoded by SEQ ID No:6; and a human *Ihh* polypeptide is encoded by SEQ ID No:7.

Table 1
Guide to vertebrate *hedgehog* sequences

	Nucleotide	Amino Acid
Chicken <i>Shh</i>	SEQ ID No. 1	SEQ ID No. 8
Mouse <i>Dhh</i>	SEQ ID No. 2	SEQ ID No. 9
Mouse <i>Ihh</i>	SEQ ID No. 3	SEQ ID No. 10
Mouse <i>Shh</i>	SEQ ID No. 4	SEQ ID No. 11
Zebrafish <i>Shh</i>	SEQ ID No. 5	SEQ ID No. 12
Human <i>Shh</i>	SEQ ID No. 6	SEQ ID No. 13
Human <i>Ihh</i>	SEQ ID No. 7	SEQ ID No. 14

Certain of the vertebrate *Hedgehog* proteins (*hh*) of the present invention are defined by SEQ ID Nos:8-14 and can be cloned from vertebrate organisms including fish, avian and mammalian sources. These proteins are distinct from the *Drosophila* protein referred to in the literature as a hedgehog protein which, for clarity, will be referred to hereinafter as "Dros-HH". In addition to the sequence variation between the various *hh* homologs, the vertebrate *hedgehog* proteins are apparently present naturally in a number of different forms, including a pro-form, a full-length mature form, and several processed fragments thereof. The pro-form includes an N-terminal signal peptide for directed secretion of the extracellular domain, while the full-length mature form lacks this signal sequence. Further processing of the mature form apparently occurs in some instances to yield biologically active fragments of the protein. For instance, *sonic hedgehog* undergoes additional proteolytic processing to yield two peptides of approximately 19 kDa and 27 kDa, both of which are secreted. In addition to proteolytic fragmentation, the vertebrate *hedgehog* proteins can also be modified post-translationally, such as by glycosylation, though bacterially produced (e.g. unglycosylated) forms of the proteins apparently still maintain at least some of the activity of the native protein.

As described in the following examples, the cDNA clones provided by the present invention were first obtained by screening a mouse genomic library with a partial *Drosophila hh* cDNA clone (.7kb). Positive plaques were identified and one mouse clone was selected. This clone was then used as a probe to obtain a genomic clone containing the full coding sequence of the Mouse *Dhh* gene. As described in the attached Examples, Northern blots and

in situ hybridization demonstrated that Mouse *Dhh* is expressed in the testes, and potentially the ovaries, and is also associated with sensory neurons of the head and trunk. *Dhh* is clearly a secreted factor expressed by Sertoli cells in the male testes, which is required for maintenance of the male germ line as probably a mitotic and survival factor. *Dhh* mutants are male sterile. Furthermore, *Dhh* is expressed as one of the first signs of differentiation of the gonad, thus *Dhh* may be a target of the sex determining gene, *Sry*. Interestingly, no expression was detected on the nerve cell bodies themselves (only the axons), indicating that *Dhh* is likely produced by the Schwann cells.

In order to obtain cDNA clones encoding chicken *hh* genes, degenerate oligonucleotides were designed corresponding to the amino and carboxy ends of *Drosophila hh* exon 2. As described in the Examples below, these oligonucleotides were used to isolate PCR fragments from chicken genomic DNA. These fragments were then cloned and sequenced. Ten clones yielded two different *hh* homologs, chicken *Dhh* and chicken *Shh*. The chicken *Shh* clone was then used to screen a stage 21/22 limb bud cDNA library which yielded a full length *Shh* clone.

In order to identify other vertebrate *hedgehog* homologs, the chicken clones (*Dhh* and *Shh*) were used to probe a genomic southern blot containing chicken DNA. As described below, genomic DNA was cut with various enzymes which do not cleave within the probe sequences. The DNA was run on a gel and transferred to a nylon filter. Probes were derived by ligating each 220 bp clone into a concatomer and then labeling with a random primer kit. The blots were hybridized and washed at low stringency. In each case, three hybridizing bands were observed following autoradiography, one of which was significantly more intense (a different band with each probe), indicating that there are at least three vertebrate *hh* genes. Additional cDNA and genomic screens carried out have yielded clones of three *hh* homologs from chickens and mice (*Shh*, *Dhh* and *Ihh*), and four *hh* homologs from zebrafish (*Shh*, *Dhh*, *Ihh* and *Mhh*). Weaker hybridization signals suggested that the gene family may be even larger. Moreover, a number of weakly hybridizing genomic clones have been isolated. Subsequently, the same probes derived from chicken *hedgehog* homologs have been utilized to screen a human genomic library. PCR fragments derived from the human genomic library were then sequenced, and PCR probes derived from the human sequences were used to screen human fetal cDNA libraries. Full-length cDNA encoding human sonic *hedgehog* protein (*Shh*) and partial cDNA encoding human Indian *hedgehog* protein (*Ihh*) were isolated from the fetal library, and represent a source of recombinant human *hedgehog* proteins.

To order to determine the expression patterns of the various vertebrate *hh* homologs, *in situ* hybridizations were performed in developing embryos of chicken, mice and fish. As described in the Examples below, the resulting expression patterns of each *hh* homolog were similar across each species and revealed that *hh* genes are expressed in a number of important

embryonic signaling centers. For example, *Shh* is expressed in Hensen's node, the notochord, the ventral floorplate of the developing neural tube, and the ZPA at the base of the limb buds. *Shh* is also expressed in differentiated motor neurons in the embryonic mouse (at 11.5 days post fertilization), therefore, *Shh* may play a role in later stages of motor neuron development, perhaps in proliferation, but more likely in survival of this cell population. *Ihh* is expressed in the embryonic yolk sac and hindgut, and appear also to be involved in chondrogenesis; *Dhh* is expressed in the testes; and *Mhh* (only in zebrafish) is expressed in the notochord and in certain cranial nerves.

Furthermore, experimental evidence indicates that certain *hedgehog* proteins initiate expression of secondary signaling molecules, including *Bmp-2* (a TGF- β relative) in the mesoderm and *Fgf-4* in the ectoderm. The mesoderm requires ectodermally-derived competence factor(s), which include *Fgf-4*, to activate target gene expression in response to *hedgehog* signaling. The expression of, for example, Sonic and *Fgf-4* is coordinately regulated by a positive feedback loop operating between the posterior mesoderm and the overlying AER, which is the ridge of pseudostratified epithelium extending antero-posteriorly along the distal margin of the bud. These data provide a basis for understanding the integration of growth and patterning in the developing limb which can have important implications in the treatment of bone disorders described in greater detail herein.

To determine the role *hedgehog* proteins plays in inductive interactions between the endoderm and mesoderm, which are critical to gut morphogenesis, in situ hybridizations and recombinant retroviral injections were performed in developing chick embryos. The ventral mesoderm is induced to undergo gut-specific differentiation by the adjacent endoderm. As described in Examples below, at the earliest stages of chick gut formation *Shh* is expressed by the endoderm, and *BMP-4* (a TGF- β relative) is expressed in the adjacent visceral mesoderm. Ectopic expression of *Sonic* is sufficient to induce expression of *BMP-4* in visceral mesoderm, suggesting that *Sonic* serves as an inductive signal from the endoderm to the mesoderm. Subsequent organ-specific endodermal differentiation depends on regional inductive signal from the visceral mesoderm. Hox genes are expressed in the undifferentiated chick hind gut mesoderm with boundaries corresponding to morphologic borders, suggesting a role in regulating gut morphogenesis.

Accordingly, certain aspects of the present invention relate to nucleic acids encoding vertebrate *hedgehog* proteins, the *hedgehog* proteins themselves, antibodies immunoreactive with *hh* proteins, and preparations of such compositions. Moreover, the present invention provides diagnostic and therapeutic assays and reagents for detecting and treating disorders involving, for example, aberrant expression of vertebrate *hedgehog* homologs. In addition, drug discovery assays are provided for identifying agents which can modulate the binding of vertebrate *hedgehog* homologues to *hedgehog*-binding moieties (such as *hedgehog* receptors,

ligands, or other extracellular matrix components). Such agents can be useful therapeutically to alter the growth and/or differentiation of a cell. Other aspects of the invention are described below or will be apparent to those skilled in the art in light of the present disclosure.

5 For convenience, certain terms employed in the specification, examples, and appended claims are collected here.

As used herein, the term "nucleic acid" refers to polynucleotides such as deoxyribonucleic acid (DNA), and, where appropriate, ribonucleic acid (RNA). The term should also be understood to include, as equivalents, analogs of either RNA or DNA made
10 from nucleotide analogs, and, as applicable to the embodiment being described, single (sense or antisense) and double-stranded polynucleotides.

As used herein, the term "gene" or "recombinant gene" refers to a nucleic acid comprising an open reading frame encoding one of the vertebrate *hh* polypeptides of the present invention, including both exon and (optionally) intron sequences. A "recombinant
15 gene" refers to nucleic acid encoding a vertebrate *hh* polypeptide and comprising vertebrate *hh*-encoding exon sequences, though it may optionally include intron sequences which are either derived from a chromosomal vertebrate *hh* gene or from an unrelated chromosomal gene. Exemplary recombinant genes encoding the subject vertebrate *hh* polypeptide are represented by SEQ ID No:1, SEQ ID No:2, SEQ ID No:3, SEQ ID No:4, SEQ ID No:5,
20 SEQ ID No:6 or SEQ ID No:7. The term "intron" refers to a DNA sequence present in a given vertebrate *hh* gene which is not translated into protein and is generally found between exons.

As used herein, the term "transfection" means the introduction of a nucleic acid, e.g., an expression vector, into a recipient cell by nucleic acid-mediated gene transfer.
25 "Transformation", as used herein, refers to a process in which a cell's genotype is changed as a result of the cellular uptake of exogenous DNA or RNA, and, for example, the transformed cell expresses a recombinant form of a vertebrate *hh* polypeptide or, where anti-sense expression occurs from the transferred gene, the expression of a naturally-occurring form of the vertebrate *hh* protein is disrupted.

30 As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of preferred vector is an episome, i.e., a nucleic acid capable of extra-chromosomal replication. Preferred vectors are those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively
35 linked are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer generally to

circular double stranded DNA loops which, in their vector form are not bound to the chromosome. In the present specification, "plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which become known in the art subsequently hereto.

"Transcriptional regulatory sequence" is a generic term used throughout the specification to refer to DNA sequences, such as initiation signals, enhancers, and promoters, which induce or control transcription of protein coding sequences with which they are operably linked. In preferred embodiments, transcription of one of the recombinant vertebrate *hedgehog* genes is under the control of a promoter sequence (or other transcriptional regulatory sequence) which controls the expression of the recombinant gene in a cell-type in which expression is intended. It will also be understood that the recombinant gene can be under the control of transcriptional regulatory sequences which are the same or which are different from those sequences which control transcription of the naturally-occurring forms of *hedgehog* proteins.

As used herein, the term "tissue-specific promoter" means a DNA sequence that serves as a promoter, i.e., regulates expression of a selected DNA sequence operably linked to the promoter, and which effects expression of the selected DNA sequence in specific cells of a tissue, such as cells of neural origin, e.g. neuronal cells. The term also covers so-called "leaky" promoters, which regulate expression of a selected DNA primarily in one tissue, but cause expression in other tissues as well.

As used herein, a "transgenic animal" is any animal, preferably a non-human mammal, bird or an amphibian, in which one or more of the cells of the animal contain heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA. In the typical transgenic animals described herein, the transgene causes cells to express a recombinant form of one of the vertebrate *hh* proteins, e.g. either agonistic or antagonistic forms. However, transgenic animals in which the recombinant vertebrate *hh* gene is silent are also contemplated, as for example, the FLP or CRE recombinase dependent constructs described below. The "non-human animals" of the invention include vertebrates such as rodents, non-human primates, sheep, dog, cow, chickens, amphibians, reptiles, etc. Preferred non-human animals are selected from the rodent family including rat and mouse, most

preferably mouse, though transgenic amphibians, such as members of the *Xenopus* genus, and transgenic chickens can also provide important tools for understanding and identifying agents which can affect, for example, embryogenesis and tissue formation. The term "chimeric animal" is used herein to refer to animals in which the recombinant gene is found, or in which
5 the recombinant is expressed in some but not all cells of the animal. The term "tissue-specific chimeric animal" indicates that one of the recombinant vertebrate *hh* genes is present and/or expressed in some tissues but not others.

As used herein, the term "transgene" means a nucleic acid sequence (encoding, e.g., one of the vertebrate *hh* polypeptides), which is partly or entirely heterologous, i.e., foreign,
10 to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout). A transgene can include one or
15 more transcriptional regulatory sequences and any other nucleic acid, such as introns, that may be necessary for optimal expression of a selected nucleic acid.

As is well known, genes for a particular polypeptide may exist in single or multiple copies within the genome of an individual. Such duplicate genes may be identical or may have certain modifications, including nucleotide substitutions, additions or deletions, which
20 all still code for polypeptides having substantially the same activity. The term "DNA sequence encoding a vertebrate *hh* polypeptide" may thus refer to one or more genes within a particular individual. Moreover, certain differences in nucleotide sequences may exist between individual organisms, which are called alleles. Such allelic differences may or may not result in differences in amino acid sequence of the encoded polypeptide yet still encode a
25 protein with the same biological activity.

"Homology" refers to sequence similarity between two peptides or between two nucleic acid molecules. Homology can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same base or amino acid, then the molecules are
30 homologous at that position. A degree of homology between sequences is a function of the number of matching or homologous positions shared by the sequences. An "unrelated" or "non-homologous" sequence shares less than 40 percent identity, though preferably less than 25 percent identity, with one of the vertebrate *hh* sequences of the present invention.

"Cells," "host cells" or "recombinant host cells" are terms used interchangeably
35 herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in

succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A "chimeric protein" or "fusion protein" is a fusion of a first amino acid sequence encoding one of the subject vertebrate *hh* polypeptides with a second amino acid sequence defining a domain foreign to and not substantially homologous with any domain of one of the vertebrate *hh* proteins. A chimeric protein may present a foreign domain which is found (albeit in a different protein) in an organism which also expresses the first protein, or it may be an "interspecies", "intergenic", etc. fusion of protein structures expressed by different kinds of organisms. In general, a fusion protein can be represented by the general formula X-*hh*-Y, wherein *hh* represents a portion of the protein which is derived from one of the vertebrate *hh* proteins, and X and Y are independently absent or represent amino acid sequences which are not related to one of the vertebrate *hh* sequences in an organism, including naturally occurring mutants.

As used herein, the terms "transforming growth factor-beta" and "TGF- β " denote a family of structurally related paracrine polypeptides found ubiquitously in vertebrates, and prototypic of a large family of metazoan growth, differentiation, and morphogenesis factors (see, for review, Massague et al. (1990) *Ann Rev Cell Biol* 6:597-641; and Sporn et al. (1992) *J Cell Biol* 119:1017-1021). Included in this family are the "bone morphogenetic proteins" or "BMPs", which refers to proteins isolated from bone, and fragments thereof and synthetic peptides which are capable of inducing bone deposition alone or when combined with appropriate cofactors. Preparation of BMPs, such as BMP-1, -2, -3, and -4, is described in, for example, PCT publication WO 88/00205. Wozney (1989) *Growth Fact Res* 1:267-280 describes additional BMP proteins closely related to BMP-2, and which have been designated BMP-5, -6, and -7. PCT publications WO89/09787 and WO89/09788 describe a protein called "OP-1," now known to be BMP-7. Other BMPs are known in the art.

The term "isolated" as also used herein with respect to nucleic acids, such as DNA or RNA, refers to molecules separated from other DNAs, or RNAs, respectively, that are present in the natural source of the macromolecule. For example, an isolated nucleic acid encoding one of the subject vertebrate *hh* polypeptides preferably includes no more than 10 kilobases (kb) of nucleic acid sequence which naturally immediately flanks the vertebrate *hh* gene in genomic DNA, more preferably no more than 5kb of such naturally occurring flanking sequences, and most preferably less than 1.5kb of such naturally occurring flanking sequence. The term isolated as used herein also refers to a nucleic acid or peptide that is substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized.

20

Moreover, an "isolated nucleic acid" is meant to include nucleic acid fragments which are not naturally occurring as fragments and would not be found in the natural state.

As described below, one aspect of the invention pertains to isolated nucleic acids comprising the nucleotide sequences encoding vertebrate *hh* homologues, and/or equivalents of such nucleic acids. The term nucleic acid as used herein is intended to include fragments as equivalents. The term equivalent is understood to include nucleotide sequences encoding functionally equivalent *hh* polypeptides or functionally equivalent peptides having an activity of a vertebrate *hh* protein such as described herein. Equivalent nucleotide sequences will include sequences that differ by one or more nucleotide substitutions, additions or deletions, such as allelic variants; and will, therefore, include sequences that differ from the nucleotide sequence of the vertebrate *hh* cDNAs shown in SEQ ID Nos:1-7 due to the degeneracy of the genetic code. Equivalents will also include nucleotide sequences that hybridize under stringent conditions (i.e., equivalent to about 20-27°C below the melting temperature (T_m) of the DNA duplex formed in about 1M salt) to the nucleotide sequences represented in SEQ ID Nos:1-7. In one embodiment, equivalents will further include nucleic acid sequences derived from and evolutionarily related to, a nucleotide sequences shown in any of SEQ ID Nos:1-7.

Moreover, it will be generally appreciated that, under certain circumstances, it may be advantageous to provide homologs of one of the subject *hedgehog* polypeptides which function in a limited capacity as one of either an *hh* agonist or an *hh* antagonist, in order to promote or inhibit only a subset of the biological activities of the naturally-occurring form of the protein. Thus, specific biological effects can be elicited by treatment with a homolog of limited function, and with fewer side effects relative to treatment with agonists or antagonists which are directed to all of the biological activities of naturally occurring forms of *hedgehog* proteins.

Homologs of one of the subject *hedgehog* proteins can be generated by mutagenesis, such as by discrete point mutation(s), or by truncation. For instance, mutation can give rise to homologs which retain substantially the same, or merely a subset, of the biological activity of the *hh* polypeptide from which it was derived. Alternatively, antagonistic forms of the protein can be generated which are able to inhibit the function of the naturally occurring form of the protein, such as by competitively binding to an *hh* receptor.

Polypeptides referred to herein as having an activity of a vertebrate *hh* protein are defined as peptides that have an amino acid sequence corresponding to all or a portion of the amino acid sequences of a vertebrate *hh* proteins shown in any of SEQ ID No:8, SEQ ID No:9, SEQ ID No:10, SEQ ID No:11, SEQ ID No:12, SEQ ID No:13 or SEQ ID No:14 and which have at least one biological activity of a vertebrate *hh* protein. Examples of such biological activity of a vertebrate *hh* protein include the ability to induce (or otherwise

modulate) formation and differentiation of the head, limbs, lungs, central nervous system (CNS), or mesodermal patterning of developing vertebrate embryos. In preferred embodiments, the biological activity can comprise an ability to regulate neurogenesis, such as a motor neuron inducing activity, a neuronal differentiation inducing activity, or a neuronal survival promoting activity. *Hedgehog* proteins of the present invention can also have biological activities which include an ability to regulate organogenesis, such as through the ability to influence limb patterning, by, for example, skeletogenic activity. The biological activity associated with the *hedgehog* proteins of the present invention can also include the ability to induce stem cell or germ cell differentiation, including the ability to induce differentiation of chondrocytes or an involvement in spermatogenesis. *Hedgehog* proteins of the present invention can also be characterized in terms of biological activities which include: an ability to modulate proliferation, survival and/or differentiation of mesodermally-derived tissue, such as tissue derived from dorsal mesoderm; the ability to modulate proliferation, survival and/or differentiation of ectodermally-derived tissue, such as tissue derived from the neural tube, neural crest, or head mesenchyme; the ability to modulate proliferation, survival and/or differentiation of endodermally-derived tissue, such as tissue derived from the primitive gut. Moreover, as described in the Examples below, the subject *hedgehog* proteins have the ability to induce expression of secondary signaling molecules, such as members of the Transforming Growth Factor β (TGF β) family, including bone morphogenic proteins, e.g. *BMP-2* and *BMP-4*, as well as members of the fibroblast growth factor (FGF) family, such as *Fgf-4*. Other biological activities of the subject *hedgehog* proteins are described herein or will be reasonably apparent to those skilled in the art. According to the present invention, a polypeptide has biological activity if it is a specific agonist or antagonist of a naturally-occurring form of a vertebrate *hedgehog* protein.

Preferred nucleic acids encode a vertebrate *hedgehog* polypeptide comprising an amino acid sequence at least 60% homologous, more preferably 70% homologous and most preferably 80% homologous with an amino acid sequence selected from the group consisting of SEQ ID Nos:8-14. Nucleic acids which encode polypeptides at least about 90%, more preferably at least about 95%, and most preferably at least about 98-99% homology with an amino acid sequence represented in one of SEQ ID Nos:8-14 are also within the scope of the invention. In one embodiment, the nucleic acid is a cDNA encoding a peptide having at least one activity of the subject vertebrate *hh* polypeptide. Preferably, the nucleic acid includes all or a portion of the nucleotide sequence corresponding to the coding region of SEQ ID Nos:1-7.

Another aspect of the invention provides a nucleic acid which hybridizes under high or low stringency conditions to a nucleic acid represented by one of SEQ ID Nos:1-7. Appropriate stringency conditions which promote DNA hybridization, for example, 6.0 x

2.2

sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 x SSC at 50°C, are known to those skilled in the art or can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0 x SSC at 50°C to a high stringency of about 0.2 x SSC at 50°C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22°C, to high stringency conditions at about 65°C.

Nucleic acids, having a sequence that differs from the nucleotide sequences shown in one of SEQ ID No:1, SEQ ID No:2, SEQ ID No:3, SEQ ID No:4, SEQ ID No:5, SEQ ID No:6 or SEQ ID No:7 due to degeneracy in the genetic code are also within the scope of the invention. Such nucleic acids encode functionally equivalent peptides (i.e., a peptide having a biological activity of a vertebrate *hh* polypeptide) but differ in sequence from the sequence shown in the sequence listing due to degeneracy in the genetic code. For example, a number of amino acids are designated by more than one triplet. Codons that specify the same amino acid, or synonyms (for example, CAU and CAC each encode histidine) may result in "silent" mutations which do not affect the amino acid sequence of a vertebrate *hh* polypeptide. However, it is expected that DNA sequence polymorphisms that do lead to changes in the amino acid sequences of the subject *hh* polypeptides will exist among vertebrates. One skilled in the art will appreciate that these variations in one or more nucleotides (up to about 3-5% of the nucleotides) of the nucleic acids encoding polypeptides having an activity of a vertebrate *hh* polypeptide may exist among individuals of a given species due to natural allelic variation.

Fragments of the nucleic acids encoding an active portion of the vertebrate *hedgehog* proteins are also within the scope of the invention. As used herein, a *hedgehog* gene fragment refers to a nucleic acid having fewer nucleotides than the nucleotide sequence encoding the entire amino acid sequence of a vertebrate *hh* protein represented in SEQ ID No:8, SEQ ID No:9, SEQ ID No:10, SEQ ID No:11, SEQ ID No:12, SEQ ID No:13 or SEQ ID No:14, yet which (preferably) encodes a peptide which retains some biological activity of the full length protein, e.g. the fragment retains the ability to induce formation and differentiation of the head, limbs, lungs, central nervous system (CNS), or mesodermal patterning of developing vertebrate embryo. Nucleic acid fragments within the scope of the present invention include those capable of hybridizing under high or low stringency conditions with nucleic acids from other species for use in screening protocols to detect other *hedgehog* homologs, as well as those capable of hybridizing with nucleic acids from human specimens for use in detecting the presence of a nucleic acid encoding a *hedgehog* protein, including alternate isoforms, e.g. mRNA splicing variants. Nucleic acids within the scope of the invention may also contain linker sequences, modified restriction endonuclease sites and

other sequences useful for molecular cloning, expression or purification of recombinant forms of the subject *hh* polypeptides.

As indicated by the examples set out below, *hedgehog* protein-encoding nucleic acids can be obtained from mRNA present in any of a number of eukaryotic cells. It should also be possible to obtain nucleic acids encoding vertebrate *hh* polypeptides of the present invention from genomic DNA obtained from both adults and embryos. For example, a gene encoding a *hh* protein can be cloned from either a cDNA or a genomic library in accordance with protocols described herein, as well as those generally known to persons skilled in the art. A cDNA encoding a *hedgehog* protein can be obtained by isolating total mRNA from a cell, e.g. a mammalian cell, e.g. a human cell, including embryonic cells. Double stranded cDNAs can then be prepared from the total mRNA, and subsequently inserted into a suitable plasmid or bacteriophage vector using any one of a number of known techniques. The gene encoding a vertebrate *hh* protein can also be cloned using established polymerase chain reaction techniques in accordance with the nucleotide sequence information provided by the invention. The nucleic acid of the invention can be DNA or RNA. A preferred nucleic acid is a cDNA represented by a sequence selected from the group consisting of SEQ ID Nos:1-7.

Another aspect of the invention relates to the use of the isolated nucleic acid in "antisense" therapy. As used herein, "antisense" therapy refers to administration or *in situ* generation of oligonucleotide probes or their derivatives which specifically hybridizes (e.g. binds) under cellular conditions, with the cellular mRNA and/or genomic DNA encoding one or more of the subject *hedgehog* proteins so as to inhibit expression of that protein, e.g. by inhibiting transcription and/or translation. The binding may be by conventional base pair complementarity, or, for example, in the case of binding to DNA duplexes, through specific interactions in the major groove of the double helix. In general, "antisense" therapy refers to the range of techniques generally employed in the art, and includes any therapy which relies on specific binding to oligonucleotide sequences.

An antisense construct of the present invention can be delivered, for example, as an expression plasmid which, when transcribed in the cell, produces RNA which is complementary to at least a unique portion of the cellular mRNA which encodes a vertebrate *hh* protein. Alternatively, the antisense construct is an oligonucleotide probe which is generated *ex vivo* and which, when introduced into the cell causes inhibition of expression by hybridizing with the mRNA and/or genomic sequences of a vertebrate *hh* gene. Such oligonucleotide probes are preferably modified oligonucleotide which are resistant to endogenous nucleases, e.g. exonucleases and/or endonucleases, and is therefore stable *in vivo*. Exemplary nucleic acid molecules for use as antisense oligonucleotides are phosphoramidate, phosphothioate and methylphosphonate analogs of DNA (see also U.S. Patents 5,176,996; 5,264,564; and 5,256,775). Additionally, general approaches to

24

constructing oligomers useful in antisense therapy have been reviewed, for example, by Van der Krol et al. (1988) *Biotechniques* 6:958-976; and Stein et al. (1988) *Cancer Res* 48:2659-2668.

Accordingly, the modified oligomers of the invention are useful in therapeutic, diagnostic, and research contexts. In therapeutic applications, the oligomers are utilized in a manner appropriate for antisense therapy in general. For such therapy, the oligomers of the invention can be formulated for a variety of loads of administration, including systemic and topical or localized administration. Techniques and formulations generally may be found in Remington's Pharmaceutical Sciences, Meade Publishing Co., Easton, PA. For systemic administration, injection is preferred, including intramuscular, intravenous, intraperitoneal, and subcutaneous for injection, the oligomers of the invention can be formulated in liquid solutions, preferably in physiologically compatible buffers such as Hank's solution or Ringer's solution. In addition, the oligomers may be formulated in solid form and redissolved or suspended immediately prior to use. Lyophilized forms are also included.

Systemic administration can also be by transmucosal or transdermal means, or the compounds can be administered orally. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration bile salts and fusidic acid derivatives. In addition, detergents may be used to facilitate permeation. Transmucosal administration may be through nasal sprays or using suppositories. For oral administration, the oligomers are formulated into conventional oral administration forms such as capsules, tablets, and tonics. For topical administration, the oligomers of the invention are formulated into ointments, salves, gels, or creams as generally known in the art.

In addition to use in therapy, the oligomers of the invention may be used as diagnostic reagents to detect the presence or absence of the target DNA or RNA sequences to which they specifically bind. Such diagnostic tests are described in further detail below.

Likewise, the antisense constructs of the present invention, by antagonizing the normal biological activity of one of the *hedgehog* proteins, can be used in the manipulation of tissue, e.g. tissue differentiation, both *in vivo* and in *ex vivo* tissue cultures.

Also, the anti-sense techniques (e.g. microinjection of antisense molecules, or transfection with plasmids whose transcripts are anti-sense with regard to an *hh* mRNA or gene sequence) can be used to investigate role of *hh* in developmental events, as well as the normal cellular function of *hh* in adult tissue. Such techniques can be utilized in cell culture, but can also be used in the creation of transgenic animals.

This invention also provides expression vectors containing a nucleic acid encoding a vertebrate *hh* polypeptide, operably linked to at least one transcriptional regulatory sequence. Operably linked is intended to mean that the nucleotide sequence is linked to a regulatory sequence in a manner which allows expression of the nucleotide sequence. Regulatory sequences are art-recognized and are selected to direct expression of the subject vertebrate *hh* proteins. Accordingly, the term transcriptional regulatory sequence includes promoters, enhancers and other expression control elements. Such regulatory sequences are described in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). For instance, any of a wide variety of expression control sequences, sequences that control the expression of a DNA sequence when operatively linked to it, may be used in these vectors to express DNA sequences encoding vertebrate *hh* polypeptides of this invention. Such useful expression control sequences, include, for example, a viral LTR, such as the LTR of the Moloney murine leukemia virus, the early and late promoters of SV40, adenovirus or cytomegalovirus immediate early promoter, the lac system, the trp system, the TAC or TRC system, T7 promoter whose expression is directed by T7 RNA polymerase, the major operator and promoter regions of phage λ , the control regions for fd coat protein, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, e.g., Pho5, the promoters of the yeast α -mating factors, the polyhedron promoter of the baculovirus system and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof. It should be understood that the design of the expression vector may depend on such factors as the choice of the host cell to be transformed and/or the type of protein desired to be expressed. Moreover, the vector's copy number, the ability to control that copy number and the expression of any other proteins encoded by the vector, such as antibiotic markers, should also be considered. In one embodiment, the expression vector includes a recombinant gene encoding a peptide having an agonistic activity of a subject *hedgehog* polypeptide, or alternatively, encoding a peptide which is an antagonistic form of the *hh* protein. Such expression vectors can be used to transfect cells and thereby produce polypeptides, including fusion proteins, encoded by nucleic acids as described herein.

Moreover, the gene constructs of the present invention can also be used as a part of a gene therapy protocol to deliver nucleic acids encoding either an agonistic or antagonistic form of one of the subject vertebrate *hedgehog* proteins. Thus, another aspect of the invention features expression vectors for *in vivo* or *in vitro* transfection and expression of a vertebrate *hh* polypeptide in particular cell types so as to reconstitute the function of, or alternatively, abrogate the function of *hedgehog*-induced signaling in a tissue in which the naturally-occurring form of the protein is misexpressed; or to deliver a form of the protein which alters differentiation of tissue, or which inhibits neoplastic transformation.

26

Expression constructs of the subject vertebrate *hh* polypeptide, and mutants thereof, may be administered in any biologically effective carrier, e.g. any formulation or composition capable of effectively delivering the recombinant gene to cells *in vivo*. Approaches include insertion of the subject gene in viral vectors including recombinant retroviruses, adenovirus, adeno-associated virus, and herpes simplex virus-1, or recombinant bacterial or eukaryotic plasmids. Viral vectors transfect cells directly; plasmid DNA can be delivered with the help of, for example, cationic liposomes (lipofectin) or derivatized (e.g. antibody conjugated), polylysine conjugates, gramicidin S, artificial viral envelopes or other such intracellular carriers, as well as direct injection of the gene construct or CaPO_4 precipitation carried out *in vivo*. It will be appreciated that because transduction of appropriate target cells represents the critical first step in gene therapy, choice of the particular gene delivery system will depend on such factors as the phenotype of the intended target and the route of administration, e.g. locally or systemically. Furthermore, it will be recognized that the particular gene construct provided for *in vivo* transduction of *hedgehog* expression are also useful for *in vitro* transduction of cells, such as for use in the *ex vivo* tissue culture systems described below.

A preferred approach for *in vivo* introduction of nucleic acid into a cell is by use of a viral vector containing nucleic acid, e.g. a cDNA, encoding the particular form of the *hedgehog* polypeptide desired. Infection of cells with a viral vector has the advantage that a large proportion of the targeted cells can receive the nucleic acid. Additionally, molecules encoded within the viral vector, e.g., by a cDNA contained in the viral vector, are expressed efficiently in cells which have taken up viral vector nucleic acid.

Retrovirus vectors and adeno-associated virus vectors are generally understood to be the recombinant gene delivery system of choice for the transfer of exogenous genes *in vivo*, particularly into humans. These vectors provide efficient delivery of genes into cells, and the transferred nucleic acids are stably integrated into the chromosomal DNA of the host. A major prerequisite for the use of retroviruses is to ensure the safety of their use, particularly with regard to the possibility of the spread of wild-type virus in the cell population. The development of specialized cell lines (termed "packaging cells") which produce only replication-defective retroviruses has increased the utility of retroviruses for gene therapy, and defective retroviruses are well characterized for use in gene transfer for gene therapy purposes (for a review see Miller, A.D. (1990) *Blood* 76:271). Thus, recombinant retrovirus can be constructed in which part of the retroviral coding sequence (*gag*, *pol*, *env*) has been replaced by nucleic acid encoding one of the subject proteins rendering the retrovirus replication defective. The replication defective retrovirus is then packaged into virions which can be used to infect a target cell through the use of a helper virus by standard techniques. Protocols for producing recombinant retroviruses and for infecting cells *in vitro* or *in vivo*

with such viruses can be found in Current Protocols in Molecular Biology, Ausubel, F.M. et al. (eds.) Greene Publishing Associates, (1989), Sections 9.10-9.14 and other standard laboratory manuals. Examples of suitable retroviruses include pLJ, pZIP, pWE and pEM which are well known to those skilled in the art. Examples of suitable packaging virus lines for preparing both ecotropic and amphotropic retroviral systems include ψ Crip, ψ Cre, ψ 2 and ψ Am. Retroviruses have been used to introduce a variety of genes into many different cell types, including neuronal cells, *in vitro* and/or *in vivo* (see for example Eglitis, et al. (1985) *Science* 230:1395-1398; Danos and Mulligan (1988) *Proc. Natl. Acad. Sci. USA* 85:6460-6464; Wilson et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:3014-3018; Armentano et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:6141-6145; Huber et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:8039-8043; Ferry et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:8377-8381; Chowdhury et al. (1991) *Science* 254:1802-1805; van Beusechem et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:7640-7644; Kay et al. (1992) *Human Gene Therapy* 3:641-647; Dai et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:10892-10895; Hwu et al. (1993) *J. Immunol.* 150:4104-4115; U.S. Patent No. 4,868,116; U.S. Patent No. 4,980,286; PCT Application WO 89/07136; PCT Application WO 89/02468; PCT Application WO 89/05345; and PCT Application WO 92/07573).

Furthermore, it has been shown that it is possible to limit the infection spectrum of retroviruses and consequently of retroviral-based vectors, by modifying the viral packaging proteins on the surface of the viral particle (see, for example PCT publications WO93/25234 and WO94/06920). For instance, strategies for the modification of the infection spectrum of retroviral vectors include: coupling antibodies specific for cell surface antigens to the viral *env* protein (Roux et al. (1989) *PNAS* 86:9079-9083; Julan et al. (1992) *J. Gen Virol* 73:3251-3255; and Goud et al. (1983) *Virology* 163:251-254); or coupling cell surface receptor ligands to the viral *env* proteins (Neda et al. (1991) *J Biol Chem* 266:14143-14146). Coupling can be in the form of the chemical cross-linking with a protein or other variety (e.g. lactose to convert the *env* protein to an asialoglycoprotein), as well as by generating fusion proteins (e.g. single-chain antibody/*env* fusion proteins). This technique, while useful to limit or otherwise direct the infection to certain tissue types, can also be used to convert an ecotropic vector in to an amphotropic vector.

Moreover, use of retroviral gene delivery can be further enhanced by the use of tissue- or cell-specific transcriptional regulatory sequences which control expression of the *hh* gene of the retroviral vector.

Another viral gene delivery system useful in the present invention utilizes adenovirus-derived vectors. The genome of an adenovirus can be manipulated such that it encodes and expresses a gene product of interest but is inactivated in terms of its ability to replicate in a normal lytic viral life cycle. See for example Berkner et al. (1988) *BioTechniques* 6:616;

Rosenfeld et al. (1991) *Science* 252:431-434; and Rosenfeld et al. (1992) *Cell* 68:143-155. Suitable adenoviral vectors derived from the adenovirus strain Ad type 5 dl324 or other strains of adenovirus (e.g., Ad2, Ad3, Ad7 etc.) are well known to those skilled in the art. Recombinant adenoviruses can be advantageous in certain circumstances in that they can be used to infect a wide variety of cell types, including airway epithelium (Rosenfeld et al. (1992) cited *supra*), endothelial cells (Lemarchand et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:6482-6486), hepatocytes (Herz and Gerard (1993) *Proc. Natl. Acad. Sci. USA* 90:2812-2816) and muscle cells (Quantin et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:2581-2584). Furthermore, the virus particle is relatively stable and amenable to purification and concentration, and as above, can be modified so as to affect the spectrum of infectivity. Additionally, introduced adenoviral DNA (and foreign DNA contained therein) is not integrated into the genome of a host cell but remains episomal, thereby avoiding potential problems that can occur as a result of insertional mutagenesis in situations where introduced DNA becomes integrated into the host genome (e.g., retroviral DNA). Moreover, the carrying capacity of the adenoviral genome for foreign DNA is large (up to 8 kilobases) relative to other gene delivery vectors (Berkner et al. cited *supra*; Haj-Ahmand and Graham (1986) *J. Virol.* 57:267). Most replication-defective adenoviral vectors currently in use and therefore favored by the present invention are deleted for all or parts of the viral E1 and E3 genes but retain as much as 80% of the adenoviral genetic material (see, e.g., Jones et al. (1979) *Cell* 16:683; Berkner et al., *supra*; and Graham et al. in Methods in Molecular Biology, E.J. Murray, Ed. (Humana, Clifton, NJ, 1991) vol. 7. pp. 109-127). Expression of the inserted *hedgehog* gene can be under control of, for example, the E1A promoter, the major late promoter (MLP) and associated leader sequences, the E3 promoter, or exogenously added promoter sequences.

Yet another viral vector system useful for delivery of one of the subject vertebrate *hh* genes is the adeno-associated virus (AAV). Adeno-associated virus is a naturally occurring defective virus that requires another virus, such as an adenovirus or a herpes virus, as a helper virus for efficient replication and a productive life cycle. (For a review see Muzyczka et al. *Curr. Topics in Micro. and Immunol.* (1992) 158:97-129). It is also one of the few viruses that may integrate its DNA into non-dividing cells, and exhibits a high frequency of stable integration (see for example Flotte et al. (1992) *Am. J. Respir. Cell. Mol. Biol.* 7:349-356; Samulski et al. (1989) *J. Virol.* 63:3822-3828; and McLaughlin et al. (1989) *J. Virol.* 62:1963-1973). Vectors containing as little as 300 base pairs of AAV can be packaged and can integrate. Space for exogenous DNA is limited to about 4.5 kb. An AAV vector such as that described in Tratschin et al. (1985) *Mol. Cell. Biol.* 5:3251-3260 can be used to introduce DNA into cells. A variety of nucleic acids have been introduced into different cell types using AAV vectors (see for example Hermonat et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:6466-6470; Tratschin et al. (1985) *Mol. Cell. Biol.* 4:2072-2081; Wondisford et al. (1988)

Mol. Endocrinol. 2:32-39; Tratschin et al. (1984) *J. Virol.* 51:611-619; and Flotte et al. (1993) *J. Biol. Chem.* 268:3781-3790).

In addition to viral transfer methods, such as those illustrated above, non-viral methods can also be employed to cause expression of a subject *hedgehog* polypeptide in the tissue of an animal. Most nonviral methods of gene transfer rely on normal mechanisms used by mammalian cells for the uptake and intracellular transport of macromolecules. In preferred embodiments, non-viral gene delivery systems of the present invention rely on endocytic pathways for the uptake of the subject *hh* polypeptide gene by the targeted cell. Exemplary gene delivery systems of this type include liposomal derived systems, poly-lysine conjugates, and artificial viral envelopes.

In clinical settings, the gene delivery systems for the therapeutic *hedgehog* gene can be introduced into a patient by any of a number of methods, each of which is familiar in the art. For instance, a pharmaceutical preparation of the gene delivery system can be introduced systemically, e.g. by intravenous injection, and specific transduction of the protein in the target cells occurs predominantly from specificity of transfection provided by the gene delivery vehicle, cell-type or tissue-type expression due to the transcriptional regulatory sequences controlling expression of the receptor gene, or a combination thereof. In other embodiments, initial delivery of the recombinant gene is more limited with introduction into the animal being quite localized. For example, the gene delivery vehicle can be introduced by catheter (see U.S. Patent 5,328,470) or by stereotactic injection (e.g. Chen et al. (1994) *PNAS* 91: 3054-3057). A vertebrate *hh* gene, such as any one of the clones represented in the group consisting of SEQ ID NO:1-7, can be delivered in a gene therapy construct by electroporation using techniques described, for example, by Dev et al. ((1994) *Cancer Treat Rev* 20:105-115).

The pharmaceutical preparation of the gene therapy construct can consist essentially of the gene delivery system in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery system can be produced intact from recombinant cells, e.g. retroviral vectors, the pharmaceutical preparation can comprise one or more cells which produce the gene delivery system.

Another aspect of the present invention concerns recombinant forms of the *hedgehog* proteins. Recombinant polypeptides preferred by the present invention, in addition to native *hedgehog* proteins, are at least 60% homologous, more preferably 70% homologous and most preferably 80% homologous with an amino acid sequence represented by any of SEQ ID Nos:8-14. Polypeptides which possess an activity of a *hedgehog* protein (i.e. either agonistic or antagonistic), and which are at least 90%, more preferably at least 95%, and most

30

preferably at least about 98-99% homologous with a sequence selected from the group consisting of SEQ ID Nos:8-14 are also within the scope of the invention.

The term "recombinant protein" refers to a polypeptide of the present invention which is produced by recombinant DNA techniques, wherein generally, DNA encoding a vertebrate *hh* polypeptide is inserted into a suitable expression vector which is in turn used to transform a host cell to produce the heterologous protein. Moreover, the phrase "derived from", with respect to a recombinant *hedgehog* gene, is meant to include within the meaning of "recombinant protein" those proteins having an amino acid sequence of a native *hedgehog* protein, or an amino acid sequence similar thereto which is generated by mutations including substitutions and deletions (including truncation) of a naturally occurring form of the protein.

The present invention further pertains to recombinant forms of one of the subject *hedgehog* polypeptides which are encoded by genes derived from a vertebrate organism, particularly a mammal (e.g. a human), and which have amino acid sequences evolutionarily related to the *hedgehog* proteins represented in SEQ ID Nos:8-14. Such recombinant *hh* polypeptides preferably are capable of functioning in one of either role of an agonist or antagonist of at least one biological activity of a wild-type ("authentic") *hedgehog* protein of the appended sequence listing. The term "evolutionarily related to", with respect to amino acid sequences of vertebrate *hedgehog* proteins, refers to both polypeptides having amino acid sequences which have arisen naturally, and also to mutational variants of vertebrate *hh* polypeptides which are derived, for example, by combinatorial mutagenesis. Such evolutionarily derived *hedgehog* proteins polypeptides preferred by the present invention are at least 60% homologous, more preferably 70% homologous and most preferably 80% homologous with the amino acid sequence selected from the group consisting of SEQ ID Nos:8-14. Polypeptides having at least about 90%, more preferably at least about 95%, and most preferably at least about 98-99% homology with a sequence selected from the group consisting of SEQ ID Nos:8-14 are also within the scope of the invention.

The present invention further pertains to methods of producing the subject *hedgehog* polypeptides. For example, a host cell transfected with a nucleic acid vector directing expression of a nucleotide sequence encoding the subject polypeptides can be cultured under appropriate conditions to allow expression of the peptide to occur. The polypeptide *hedgehog* may be secreted and isolated from a mixture of cells and medium containing the recombinant vertebrate *hh* polypeptide. Alternatively, the peptide may be retained cytoplasmically by removing the signal peptide sequence from the recombinant *hh* gene and the cells harvested, lysed and the protein isolated. A cell culture includes host cells, media and other byproducts. Suitable media for cell culture are well known in the art. The recombinant *hh* polypeptide can be isolated from cell culture medium, host cells, or both using techniques known in the art for purifying proteins including ion-exchange

chromatography, gel filtration chromatography, ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for such peptide. In a preferred embodiment, the recombinant *hh* polypeptide is a fusion protein containing a domain which facilitates its purification, such as an *hh*/GST fusion protein.

5 This invention also pertains to a host cell transfected to express a recombinant form of the subject *hedgehog* polypeptides. The host cell may be any prokaryotic or eukaryotic cell. Thus, a nucleotide sequence derived from the cloning of vertebrate *hedgehog* proteins, encoding all or a selected portion of the full-length protein, can be used to produce a recombinant form of a vertebrate *hh* polypeptide via microbial or eukaryotic cellular
10 processes. Ligating the polynucleotide sequence into a gene construct, such as an expression vector, and transforming or transfecting into hosts, either eukaryotic (yeast, avian, insect or mammalian) or prokaryotic (bacterial cells), are standard procedures used in producing other well-known proteins, e.g. insulin, interferons, human growth hormone, IL-1, IL-2, and the like. Similar procedures, or modifications thereof, can be employed to prepare recombinant
15 *hedgehog* polypeptides by microbial means or tissue-culture technology in accord with the subject invention.

The recombinant *hedgehog* genes can be produced by ligating nucleic acid encoding an *hh* protein, or a portion thereof, into a vector suitable for expression in either prokaryotic cells, eukaryotic cells, or both. Expression vectors for production of recombinant forms of
20 the subject *hh* polypeptides include plasmids and other vectors. For instance, suitable vectors for the expression of a *hedgehog* polypeptide include plasmids of the types: pBR322-derived plasmids, pEMBL-derived plasmids, pEX-derived plasmids, pBTac-derived plasmids and pUC-derived plasmids for expression in prokaryotic cells, such as *E. coli*.

A number of vectors exist for the expression of recombinant proteins in yeast. For
25 instance, YEP24, YIP5, YEP51, YEP52, pYES2, and YRP17 are cloning and expression vehicles useful in the introduction of genetic constructs into *S. cerevisiae* (see, for example, Broach *et al.* (1983) in *Experimental Manipulation of Gene Expression*, ed. M. Inouye Academic Press, p. 83, incorporated by reference herein). These vectors can replicate in *E. coli* due the presence of the pBR322 ori, and in *S. cerevisiae* due to the replication
30 determinant of the yeast 2 micron plasmid. In addition, drug resistance markers such as ampicillin can be used. In an illustrative embodiment, an *hh* polypeptide is produced recombinantly utilizing an expression vector generated by sub-cloning the coding sequence of one of the *hedgehog* genes represented in SEQ ID Nos:1-7.

The preferred mammalian expression vectors contain both prokaryotic sequences, to
35 facilitate the propagation of the vector in bacteria, and one or more eukaryotic transcription units that are expressed in eukaryotic cells. The pcDNAI/amp, pcDNAI/neo, pRc/CMV,

pSV2gpt, pSV2neo, pSV2-dhfr, pTk2, pRSVneo, pMSG, pSVT7, pko-neo and pHyg derived vectors are examples of mammalian expression vectors suitable for transfection of eukaryotic cells. Some of these vectors are modified with sequences from bacterial plasmids, such as pBR322, to facilitate replication and drug resistance selection in both prokaryotic and eukaryotic cells. Alternatively, derivatives of viruses such as the bovine papillomavirus (BPV-1), or Epstein-Barr virus (pHEBo, pREP-derived and p205) can be used for transient expression of proteins in eukaryotic cells. The various methods employed in the preparation of the plasmids and transformation of host organisms are well known in the art. For other suitable expression systems for both prokaryotic and eukaryotic cells, as well as general recombinant procedures, see *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989) Chapters 16 and 17.

In some instances, it may be desirable to express the recombinant *hedgehog* polypeptide by the use of a baculovirus expression system. Examples of such baculovirus expression systems include pVL-derived vectors (such as pVL1392, pVL1393 and pVL941), pAcUW-derived vectors (such as pAcUW1), and pBlueBac-derived vectors (such as the β -gal containing pBlueBac III).

When it is desirable to express only a portion of an *hh* protein, such as a form lacking a portion of the N-terminus, i.e. a truncation mutant which lacks the signal peptide, it may be necessary to add a start codon (ATG) to the oligonucleotide fragment containing the desired sequence to be expressed. It is well known in the art that a methionine at the N-terminal position can be enzymatically cleaved by the use of the enzyme methionine aminopeptidase (MAP). MAP has been cloned from *E. coli* (Ben-Bassat et al. (1987) *J. Bacteriol.* 169:751-757) and *Salmonella typhimurium* and its *in vitro* activity has been demonstrated on recombinant proteins (Miller et al. (1987) *PNAS* 84:2718-1722). Therefore, removal of an N-terminal methionine, if desired, can be achieved either *in vivo* by expressing *hedgehog*-derived polypeptides in a host which produces MAP (e.g., *E. coli* or CM89 or *S. cerevisiae*), or *in vitro* by use of purified MAP (e.g., procedure of Miller et al., *supra*).

Alternatively, the coding sequences for the polypeptide can be incorporated as a part of a fusion gene including a nucleotide sequence encoding a different polypeptide. This type of expression system can be useful under conditions where it is desirable to produce an immunogenic fragment of a *hedgehog* protein. For example, the VP6 capsid protein of rotavirus can be used as an immunologic carrier protein for portions of the *hh* polypeptide, either in the monomeric form or in the form of a viral particle. The nucleic acid sequences corresponding to the portion of a subject *hedgehog* protein to which antibodies are to be raised can be incorporated into a fusion gene construct which includes coding sequences for a late vaccinia virus structural protein to produce a set of recombinant viruses expressing

fusion proteins comprising *hh* epitopes as part of the virion. It has been demonstrated with the use of immunogenic fusion proteins utilizing the Hepatitis B surface antigen fusion proteins that recombinant Hepatitis B virions can be utilized in this role as well. Similarly, chimeric constructs coding for fusion proteins containing a portion of an *hh* protein and the poliovirus capsid protein can be created to enhance immunogenicity of the set of polypeptide antigens (see, for example, EP Publication No: 0259149; and Evans et al. (1989) *Nature* 339:385; Huang et al. (1988) *J. Virol.* 62:3855; and Schlienger et al. (1992) *J. Virol.* 66:2).

The Multiple Antigen Peptide system for peptide-based immunization can also be utilized to generate an immunogen, wherein a desired portion of an *hh* polypeptide is obtained directly from organo-chemical synthesis of the peptide onto an oligomeric branching lysine core (see, for example, Posnett et al. (1988) *JBC* 263:1719 and Nardelli et al. (1992) *J. Immunol.* 148:914). Antigenic determinants of *hh* proteins can also be expressed and presented by bacterial cells.

In addition to utilizing fusion proteins to enhance immunogenicity, it is widely appreciated that fusion proteins can also facilitate the expression of proteins, and accordingly, can be used in the expression of the vertebrate *hh* polypeptides of the present invention. For example, *hedgehog* polypeptides can be generated as glutathione-S-transferase (GST-fusion) proteins. Such GST-fusion proteins can enable easy purification of the *hedgehog* polypeptide, as for example by the use of glutathione-derivatized matrices (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al. (N.Y.: John Wiley & Sons, 1991)). In another embodiment, a fusion gene coding for a purification leader sequence, such as a poly-(His)/enterokinase cleavage site sequence, can be used to replace the signal sequence which naturally occurs at the N-terminus of the *hh* protein (e.g., of the pro-form, in order to permit purification of the poly(His)-*hh* protein by affinity chromatography using a Ni^{2+} metal resin. The purification leader sequence can then be subsequently removed by treatment with enterokinase (e.g., see Hochuli et al. (1987) *J. Chromatography* 411:177; and Janknecht et al. *PNAS* 88:8972).

Techniques for making fusion genes are known to those skilled in the art. Essentially, the joining of various DNA fragments coding for different polypeptide sequences is performed in accordance with conventional techniques, employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently

34

be annealed to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al. John Wiley & Sons: 1992).

Hedgehog polypeptides may also be chemically modified to create *hh* derivatives by forming covalent or aggregate conjugates with other chemical moieties, such as glycosyl groups, lipids, phosphate, acetyl groups and the like. Covalent derivatives of *hedgehog* proteins can be prepared by linking the chemical moieties to functional groups on amino acid sidechains of the protein or at the N-terminus or at the C-terminus of the polypeptide.

For instance, *hedgehog* proteins can be generated to include a moiety, other than sequence naturally associated with the protein, that binds a component of the extracellular matrix and enhances localization of the analog to cell surfaces. For example, sequences derived from the fibronectin "type-III repeat", such as a tetrapeptide sequence R-G-D-S (Pierschbacher et al. (1984) *Nature* 309:30-3; and Kornblihtt et al. (1985) *EMBO* 4:1755-9) can be added to the *hh* polypeptide to support attachment of the chimeric molecule to a cell through binding ECM components (Ruoslahti et al. (1987) *Science* 238:491-497; Pierschbacher et al. (1987) *J. Biol. Chem.* 262:17294-8.; Hynes (1987) *Cell* 48:549-54; and Hynes (1992) *Cell* 69:11-25).

The present invention also makes available isolated *hedgehog* polypeptides which are isolated from, or otherwise substantially free of other cellular and extracellular proteins, especially morphogenic proteins or other extracellular or cell surface associated proteins which may normally be associated with the *hedgehog* polypeptide. The term "substantially free of other cellular or extracellular proteins" (also referred to herein as "contaminating proteins") or "substantially pure or purified preparations" are defined as encompassing preparations of *hh* polypeptides having less than 20% (by dry weight) contaminating protein, and preferably having less than 5% contaminating protein. Functional forms of the subject polypeptides can be prepared, for the first time, as purified preparations by using a cloned gene as described herein. By "purified", it is meant, when referring to a peptide or DNA or RNA sequence, that the indicated molecule is present in the substantial absence of other biological macromolecules, such as other proteins. The term "purified" as used herein preferably means at least 80% by dry weight, more preferably in the range of 95-99% by weight, and most preferably at least 99.8% by weight, of biological macromolecules of the same type present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 5000, can be present). The term "pure" as used herein preferably has the same numerical limits as "purified" immediately above. "Isolated" and "purified" do not encompass either natural materials in their native state or natural materials that have been separated into components (e.g., in an acrylamide gel) but not obtained either as pure (e.g. lacking contaminating proteins, or chromatography reagents such as denaturing agents and polymers, e.g. acrylamide or agarose) substances or solutions. In preferred

embodiments, purified *hedgehog* preparations will lack any contaminating proteins from the same animal from that *hedgehog* is normally produced, as can be accomplished by recombinant expression of, for example, a human *hedgehog* protein in a non-human cell.

As described above for recombinant polypeptides, isolated *hh* polypeptides can include all or a portion of the amino acid sequences represented in SEQ ID No:8, SEQ ID No:9, SEQ ID No:10, SEQ ID No:11, SEQ ID No:12, SEQ ID No:13 or SEQ ID No:14, or a homologous sequence thereto. Preferred fragments of the subject *hedgehog* proteins correspond to the N-terminal and C-terminal proteolytic fragments of the mature protein (see, for instance, Examples 6 and 9).

Isolated peptidyl portions of *hedgehog* proteins can be obtained by screening peptides recombinantly produced from the corresponding fragment of the nucleic acid encoding such peptides. In addition, fragments can be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, a *hedgehog* polypeptide of the present invention may be arbitrarily divided into fragments of desired length with no overlap of the fragments, or preferably divided into overlapping fragments of a desired length. The fragments can be produced (recombinantly or by chemical synthesis) and tested to identify those peptidyl fragments which can function as either agonists or antagonists of a wild-type (e.g., "authentic") *hedgehog* protein.

The recombinant *hedgehog* polypeptides of the present invention also include homologs of the authentic *hedgehog* proteins, such as versions of those protein which are resistant to proteolytic cleavage, as for example, due to mutations which alter potential cleavage sequences or which inactivate an enzymatic activity associated with the protein. *Hedgehog* homologs of the present invention also include proteins which have been post-translationally modified in a manner different than the authentic protein. Exemplary derivatives of vertebrate *hedgehog* proteins include polypeptides which lack N-glycosylation sites (e.g. to produce an unglycosylated protein), or which lack N-terminal and/or C-terminal sequences.

Modification of the structure of the subject vertebrate *hh* polypeptides can be for such purposes as enhancing therapeutic or prophylactic efficacy, or stability (e.g., *ex vivo* shelf life and resistance to proteolytic degradation *in vivo*). Such modified peptides, when designed to retain at least one activity of the naturally-occurring form of the protein, are considered functional equivalents of the *hedgehog* polypeptides described in more detail herein. Such modified peptides can be produced, for instance, by amino acid substitution, deletion, or addition.

For example, it is reasonable to expect that an isolated replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar

36

replacement of an amino acid with a structurally related amino acid (i.e. isosteric and/or isoelectric mutations) will not have a major effect on the biological activity of the resulting molecule. Conservative replacements are those that take place within a family of amino acids that are related in their side chains. Genetically encoded amino acids can be divided into four families: (1) acidic = aspartate, glutamate; (2) basic = lysine, arginine, histidine; (3) nonpolar = alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar = glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine. Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids. In similar fashion, the amino acid repertoire can be grouped as (1) acidic = aspartate, glutamate; (2) basic = lysine, arginine, histidine, (3) aliphatic = glycine, alanine, valine, leucine, isoleucine, serine, threonine, with serine and threonine optionally be grouped separately as aliphatic-hydroxyl; (4) aromatic = phenylalanine, tyrosine, tryptophan; (5) amide = asparagine, glutamine; and (6) sulfur-containing = cysteine and methionine. (see, for example, *Biochemistry*, 2nd ed., Ed. by L. Stryer, WH Freeman and Co.: 1981). Whether a change in the amino acid sequence of a peptide results in a functional *hedgehog* homolog (e.g. functional in the sense that it acts to mimic or antagonize the wild-type form) can be readily determined by assessing the ability of the variant peptide to produce a response in cells in a fashion similar to the wild-type protein, or competitively inhibit such a response. Polypeptides in which more than one replacement has taken place can readily be tested in the same manner.

This invention further contemplates a method for generating sets of combinatorial mutants of the subject *hedgehog* proteins as well as truncation mutants, and is especially useful for identifying potential variant sequences (e.g. homologs) that are functional in binding to a receptor for *hedgehog* proteins. The purpose of screening such combinatorial libraries is to generate, for example, novel *hh* homologs which can act as either agonists or antagonist, or alternatively, possess novel activities all together. To illustrate, *hedgehog* homologs can be engineered by the present method to provide more efficient binding to a cognate receptor, yet still retain at least a portion of an activity associated with *hh*. Thus, combinatorially-derived homologs can be generated to have an increased potency relative to a naturally occurring form of the protein. Likewise, *hedgehog* homologs can be generated by the present combinatorial approach to act as antagonists, in that they are able to mimic, for example, binding to other extracellular matrix components (such as receptors), yet not induce any biological response, thereby inhibiting the action of authentic *hedgehog* or *hedgehog* agonists. Moreover, manipulation of certain domains of *hh* by the present method can provide domains more suitable for use in fusion proteins, such as one that incorporates portions of other proteins which are derived from the extracellular matrix and/or which bind extracellular matrix components.

In one aspect of this method, the amino acid sequences for a population of *hedgehog* homologs or other related proteins are aligned, preferably to promote the highest homology possible. Such a population of variants can include, for example, *hh* homologs from one or more species. Amino acids which appear at each position of the aligned sequences are selected to create a degenerate set of combinatorial sequences. In a preferred embodiment, the variegated library of *hedgehog* variants is generated by combinatorial mutagenesis at the nucleic acid level, and is encoded by a variegated gene library. For instance, a mixture of synthetic oligonucleotides can be enzymatically ligated into gene sequences such that the degenerate set of potential *hh* sequences are expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g. for phage display) containing the set of *hh* sequences therein.

As illustrated in Figure 5A, to analyze the sequences of a population of variants, the amino acid sequences of interest can be aligned relative to sequence homology. The presence or absence of amino acids from an aligned sequence of a particular variant is relative to a chosen consensus length of a reference sequence, which can be real or artificial. In order to maintain the highest homology in alignment of sequences, deletions in the sequence of a variant relative to the reference sequence can be represented by an amino acid space (• or *), while insertional mutations in the variant relative to the reference sequence can be disregarded and left out of the sequence of the variant when aligned. For instance, Figure 5A includes the alignment of several cloned forms of *hh* from different species. Analysis of the alignment of the *hh* clones shown in Figure 5A can give rise to the generation of a degenerate library of polypeptides comprising potential *hh* sequences.

In an illustrative embodiment, alignment of exon 1/2 encoded sequences (e.g. the N-terminal approximately 165 residues of the mature protein) of each of the *Shh* clones produces a degenerate set of *Shh* polypeptides represented by the general formula:

C-G-P-G-R-G-X(1)-G-X(2)-R-R-H-P-K-K-L-T-P-L-A-Y-K-Q-F-I-P-N-V-A-E-K-T-L-G-A-S-G-R-Y-E-G-K-I-X(3)-R-N-S-E-R-F-K-E-L-T-P-N-Y-N-P-D-I-I-F-K-D-E-E-N-T-G-A-D-R-L-M-T-Q-R-C-K-D-K-L-N-X(4)-L-A-I-S-V-M-N-X(5)-W-P-G-V-X(6)-L-R-V-T-E-G-W-D-E-D-G-H-H-X(7)-E-E-S-L-H-Y-E-G-R-A-V-D-I-T-T-S-D-R-D-X(8)-S-K-Y-G-X(9)-L-X(10)-R-L-A-V-E-A-G-F-D-W-V-Y-Y-E-S-K-A-H-I-H-C-S-V-K-A-E (SEQ ID No: 40),

wherein each of the degenerate positions "X" can be an amino acid which occurs in that position in one of the human, mouse, chicken or zebrafish *Shh* clones, or, to expand the library, each X can also be selected from amongst amino acid residue which would be conservative substitutions for the amino acids which appear naturally in each of those positions. For instance, Xaa(1) represents Gly, Ala, Val, Leu, Ile, Phe, Tyr or Trp ; Xaa(2) represents Arg, His or Lys; Xaa(3) represents Gly, Ala, Val, Leu, Ile, Ser or Thr; Xaa(4)

represents Gly, Ala, Val, Leu, Ile, Ser or Thr; Xaa(5) represents Lys, Arg, His, Asn or Gln; Xaa(6) represents Lys, Arg or His; Xaa(7) represents Ser, Thr, Tyr, Trp or Phe; Xaa(8) represents Lys, Arg or His; Xaa(9) represents Met, Cys, Ser or Thr; and Xaa(10) represents Gly, Ala, Val, Leu, Ile, Ser or Thr. In an even more expansive library, each X can be selected from any amino acid.

In similar fashion, alignment of each of the human, mouse, chicken and zebrafish *hedgehog* clones (Figure 5B), can provide a degenerate polypeptide sequence represented by the general formula:

```

10      C-G-P-G-R-G-X(1)-X(2)-X(3)-R-R-X(4)-X(5)-X(6)-P-K-X(7)-L-X(8)-
      P-L-X(9)-Y-K-Q-F-X(10)-P-X(11)-X(12)-X(13)-E-X(14)-T-L-G-A-S-G-
      X(15)-X(16)-E-G-X(17)-X(18)-X(19)-R-X(20)-S-E-R-F-X(21)-X(22)-
      L-T-P-N-Y-N-P-D-I-I-F-K-D-E-E-N-X(23)-G-A-D-R-L-M-T-X(24)-R-C-
      K-X(25)-X(26)-X(27)-N-X(28)-L-A-I-S-V-M-N-X(29)-W-P-G-V-X(30)-
15      L-R-V-T-E-G-X(31)-D-E-D-G-H-H-X(32)-X(33)-X(34)-S-L-H-Y-E-G-R-
      A-X(35)-D-I-T-T-S-D-R-D-X(36)-X(37)-K-Y-G-X(38)-L-X(39)-R-L-A-
      V-E-A-G-F-D-W-V-Y-Y-E-S-X(40)-X(41)-H-X(42)-H-X(43)-S-V-K-X(44)-
      -X(45) (SEQ ID No: 41),

```

wherein, as above, each of the degenerate positions "X" can be an amino acid which occurs in a corresponding position in one of the wild-type clones, and may also include amino acid residue which would be conservative substitutions, or each X can be any amino acid residue. In an exemplary embodiment, Xaa(1) represents Gly, Ala, Val, Leu, Ile, Pro, Phe or Tyr; Xaa(2) represents Gly, Ala, Val, Leu or Ile; Xaa(3) represents Gly, Ala, Val, Leu, Ile, Lys, His or Arg; Xaa(4) represents Lys, Arg or His; Xaa(5) represents Phe, Trp, Tyr or an amino acid gap; Xaa(6) represents Gly, Ala, Val, Leu, Ile or an amino acid gap; Xaa(7) represents Asn, Gln, His, Arg or Lys; Xaa(8) represents Gly, Ala, Val, Leu, Ile, Ser or Thr; Xaa(9) represents Gly, Ala, Val, Leu, Ile, Ser or Thr; Xaa(10) represents Gly, Ala, Val, Leu, Ile, Ser or Thr; Xaa(11) represents Ser, Thr, Gln or Asn; Xaa(12) represents Met, Cys, Gly, Ala, Val, Leu, Ile, Ser or Thr; Xaa(13) represents Gly, Ala, Val, Leu, Ile or Pro; Xaa(14) represents Arg, His or Lys; Xaa(15) represents Gly, Ala, Val, Leu, Ile, Pro, Arg, His or Lys; Xaa(16) represents Gly, Ala, Val, Leu, Ile, Phe or Tyr; Xaa(17) represents Arg, His or Lys; Xaa(18) represents Gly, Ala, Val, Leu, Ile, Ser or Thr; Xaa(19) represents Thr or Ser; Xaa(20) represents Gly, Ala, Val, Leu, Ile, Asn or Gln; Xaa(21) represents Arg, His or Lys; Xaa(22) represents Asp or Glu; Xaa(23) represents Ser or Thr; Xaa(24) represents Glu, Asp, Gln or Asn; Xaa(25) represents Glu or Asp; Xaa(26) represents Arg, His or Lys; Xaa(27) represents Gly, Ala, Val, Leu or Ile; Xaa(28) represents Gly, Ala, Val, Leu, Ile, Thr or Ser; Xaa(29) represents Met, Cys, Gln, Asn, Arg, Lys or His; Xaa(30) represents Arg, His or Lys; Xaa(31) represents Trp, Phe, Tyr, Arg, His or Lys; Xaa(32) represents Gly, Ala, Val, Leu, Ile, Ser, Thr, Tyr or Phe; Xaa(33) represents Gln, Asn, Asp or Glu; Xaa(34) represents Asp or Glu; Xaa(35) represents Gly, Ala, Val, Leu, or Ile; Xaa(36) represents Arg, His or Lys; Xaa(37) represents Asn, Gln, Thr or Ser; Xaa(38) represents Gly, Ala, Val, Leu, Ile, Ser, Thr, Met or

39

Cys; Xaa(39) represents Gly, Ala, Val, Leu, Ile, Thr or Ser; Xaa(40) represents Arg, His or Lys; Xaa(41) represents Asn, Gln, Gly, Ala, Val, Leu or Ile; Xaa(42) represents Gly, Ala, Val, Leu or Ile; Xaa(43) represents Gly, Ala, Val, Leu, Ile, Ser, Thr or Cys; Xaa(44) represents Gly, Ala, Val, Leu, Ile, Thr or Ser; and Xaa(45) represents Asp or Glu.

5 There are many ways by which the library of potential *hh* homologs can be generated from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be carried out in an automatic DNA synthesizer, and the synthetic genes then ligated into an appropriate expression vector. The purpose of a degenerate set of genes is to provide, in one mixture, all of the sequences encoding the desired set of potential *hh* sequences. The synthesis of degenerate oligonucleotides is well known in the art (see for example, Narang, SA (1983) *Tetrahedron* 39:3; Itakura et al. (1981) *Recombinant DNA, Proc 3rd Cleveland Sympos. Macromolecules*, ed. AG Walton, Amsterdam: Elsevier pp273-289; Itakura et al. (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al. (1984) *Science* 198:1056; Ike et al. (1983) *Nucleic Acid Res.* 11:477. Such techniques have been employed 10 in the directed evolution of other proteins (see, for example, Scott et al. (1990) *Science* 249:386-390; Roberts et al. (1992) *PNAS* 89:2429-2433; Devlin et al. (1990) *Science* 249:404-406; Cwirla et al. (1990) *PNAS* 87: 6378-6382; as well as U.S. Patents Nos. 5,223,409, 5,198,346, and 5,096,815).

20 A wide range of techniques are known in the art for screening gene products of combinatorial libraries made by point mutations, and for screening cDNA libraries for gene products having a certain property. Such techniques will be generally adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of *hedgehog* homologs. The most widely used techniques for screening large gene libraries typically comprises cloning the gene library into replicable expression vectors, transforming 25 appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the illustrative assays described below are amenable to high through-put analysis as necessary to screen large numbers of degenerate *hedgehog* sequences created by combinatorial mutagenesis techniques. 30

 In one embodiment, the combinatorial library is designed to be secreted (e.g. the polypeptides of the library all include a signal sequence but no transmembrane or cytoplasmic domains), and is used to transfect a eukaryotic cell that can be co-cultured with embryonic cells. A functional *hedgehog* protein secreted by the cells expressing the 35 combinatorial library will diffuse to neighboring embryonic cells and induce a particular biological response, such as to illustrate, neuronal differentiation. Using antibodies directed to epitopes of particular neuronal cells (e.g. Islet-1 or Pax-1), the pattern of detection of

neuronal induction will resemble a gradient function, and will allow the isolation (generally after several repetitive rounds of selection) of cells producing active *hedgehog* homologs. Likewise, *hh* antagonists can be selected in similar fashion by the ability of the cell producing a functional antagonist to protect neighboring cells from the effect of wild-type *hedgehog* added to the culture media.

To illustrate, target cells are cultured in 24-well microtitre plates. Other eukaryotic cells are transfected with the combinatorial *hh* gene library and cultured in cell culture inserts (e.g. Collaborative Biomedical Products, Catalog #40446) that are able to fit into the wells of the microtitre plate. The cell culture inserts are placed in the wells such that recombinant *hh* homologs secreted by the cells in the insert can diffuse through the porous bottom of the insert and contact the target cells in the microtitre plate wells. After a period of time sufficient for functional forms of a *hedgehog* protein to produce a measurable response in the target cells, the inserts are removed and the effect of the variant *hedgehog* proteins on the target cells determined. For example, where the target cell is a neural crest cell and the activity desired from the *hh* homolog is the induction of neuronal differentiation, then fluorescently-labeled antibodies specific for Islet-1 or other neuronal markers can be used to score for induction in the target cells as indicative of a functional *hh* in that well. Cells from the inserts corresponding to wells which score positive for activity can be split and re-cultured on several inserts, the process being repeated until the active clones are identified.

In yet another screening assay, the candidate *hedgehog* gene products are displayed on the surface of a cell or viral particle, and the ability of particular cells or viral particles to associate with a *hedgehog*-binding moiety (such as an *hedgehog* receptor or a ligand which binds the *hedgehog* protein) via this gene product is detected in a "panning assay". Such panning steps can be carried out on cells cultured from embryos. For instance, the gene library can be cloned into the gene for a surface membrane protein of a bacterial cell, and the resulting fusion protein detected by panning (Ladner et al., WO 88/06630; Fuchs et al. (1991) *Bio/Technology* 9:1370-1371; and Goward et al. (1992) *TIBS* 18:136-140). In a similar fashion, fluorescently labeled molecules which bind *hh* can be used to score for potentially functional *hh* homologs. Cells can be visually inspected and separated under a fluorescence microscope, or, where the morphology of the cell permits, separated by a fluorescence-activated cell sorter.

In an alternate embodiment, the gene library is expressed as a fusion protein on the surface of a viral particle. For instance, in the filamentous phage system, foreign peptide sequences can be expressed on the surface of infectious phage, thereby conferring two significant benefits. First, since these phage can be applied to affinity matrices at very high concentrations, large number of phage can be screened at one time. Second, since each infectious phage displays the combinatorial gene product on its surface, if a particular phage

41

is recovered from an affinity matrix in low yield, the phage can be amplified by another round of infection. The group of almost identical *E. coli* filamentous phages M13, fd, and f1 are most often used in phage display libraries, as either of the phage gIII or gVIII coat proteins can be used to generate fusion proteins without disrupting the ultimate packaging of the viral particle (Ladner et al. PCT publication WO 90/02909; Garrard et al., PCT publication WO 92/09690; Marks et al. (1992) J. Biol. Chem. 267:16007-16010; Griffiths et al. (1993) EMBO J 12:725-734; Clackson et al. (1991) Nature 352:624-628; and Barbas et al. (1992) PNAS 89:4457-4461).

In an illustrative embodiment, the recombinant phage antibody system (RPAS, Pharmacia Catalog number 27-9400-01) can be easily modified for use in expressing and screening *hh* combinatorial libraries. For instance, the pCANTAB 5 phagemid of the RPAS kit contains the gene which encodes the phage gIII coat protein. The *hh* combinatorial gene library can be cloned into the phagemid adjacent to the gIII signal sequence such that it will be expressed as a gIII fusion protein. After ligation, the phagemid is used to transform competent *E. coli* TG1 cells. Transformed cells are subsequently infected with M13KO7 helper phage to rescue the phagemid and its candidate *hh* gene insert. The resulting recombinant phage contain phagemid DNA encoding a specific candidate *hh*, and display one or more copies of the corresponding fusion coat protein. The phage-displayed candidate *hedgehog* proteins which are capable of binding an *hh* receptor are selected or enriched by panning. For instance, the phage library can be applied to cultured embryonic cells and unbound phage washed away from the cells. The bound phage is then isolated, and if the recombinant phage express at least one copy of the wild type gIII coat protein, they will retain their ability to infect *E. coli*. Thus, successive rounds of reinfection of *E. coli*, and panning will greatly enrich for *hh* homologs, which can then be screened for further biological activities in order to differentiate agonists and antagonists. Moreover, differential panning, e.g., with two or more different *hh*-responsive cells, can facilitate isolation of *hedgehog* homologs of selectively narrower biological activity relative to the wild-type protein.

The invention also provides for reduction of the vertebrate *hh* protein to generate mimetics, e.g. peptide or non-peptide agents, which are able to disrupt binding of a vertebrate *hh* polypeptide of the present invention with an *hh* receptor. Thus, such mutagenic techniques as described above are also useful to map the determinants of the *hedgehog* proteins which participate in protein-protein interactions involved in, for example, binding of the subject vertebrate *hh* polypeptide to other extracellular matrix components. To illustrate, the critical residues of a subject *hh* polypeptide or *hh* ligand which are involved in molecular recognition of an *hh* receptor can be determined and used to generate *hedgehog*-derived peptidomimetics which competitively inhibit binding of the authentic *hedgehog* protein with that moiety. By employing, for example, scanning mutagenesis to map the amino acid

residues of each of the subject *hedgehog* proteins which are involved in binding other extracellular proteins, peptidomimetic compounds can be generated which mimic those residues of the *hedgehog* protein which facilitate the interaction. Such mimetics may then be used to interfere with the normal function of a *hedgehog* protein. For instance, non-hydrolyzable peptide analogs of such residues can be generated using benzodiazepine (e.g., see Freidinger et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), substituted gamma lactam rings (Garvey et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), keto-methylene pseudopeptides (Ewenson et al. (1986) *J Med Chem* 29:295; and Ewenson et al. in *Peptides: Structure and Function* (Proceedings of the 9th American Peptide Symposium) Pierce Chemical Co. Rockland, IL, 1985), β -turn dipeptide cores (Nagai et al. (1985) *Tetrahedron Lett* 26:647; and Sato et al. (1986) *J Chem Soc Perkin Trans* 1:1231), and β -aminoalcohols (Gordon et al. (1985) *Biochem Biophys Res Commun* 126:419; and Dann et al. (1986) *Biochem Biophys Res Commun* 134:71).

Another aspect of the invention pertains to an antibody specifically reactive with a vertebrate *hedgehog* protein. For example, by using immunogens derived from *hedgehog* protein, e.g. based on the cDNA sequences, anti-protein/anti-peptide antisera or monoclonal antibodies can be made by standard protocols (See, for example, *Antibodies: A Laboratory Manual* ed. by Harlow and Lane (Cold Spring Harbor Press: 1988)). A mammal, such as a mouse, a hamster or rabbit can be immunized with an immunogenic form of the peptide (e.g., a vertebrate *hh* polypeptide or an antigenic fragment which is capable of eliciting an antibody response). Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other techniques well known in the art. An immunogenic portion of a *hedgehog* protein can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays can be used with the immunogen as antigen to assess the levels of antibodies. In a preferred embodiment, the subject antibodies are immunospecific for antigenic determinants of a *hedgehog* protein of a vertebrate organism, such as a mammal, e.g. antigenic determinants of a protein represented by SEQ ID Nos:8-14 or a closely related homolog (e.g. at least 85% homologous, preferably at least 90% homologous, and more preferably at least 95% homologous). In yet a further preferred embodiment of the present invention, in order to provide, for example, antibodies which are immuno-selective for discrete *hedgehog* homologs, e.g. *Shh* versus *Dhh* versus *Ihh*, the anti-*hh* polypeptide antibodies do not substantially cross react (i.e. does not react specifically) with a protein which is, for example, less than 85% homologous to any of SEQ ID Nos:8-14; e.g., less than 95% homologous with one of SEQ ID Nos:8-14; e.g., less than 98-99% homologous with

one of SEQ ID Nos:8-14. By "not substantially cross react", it is meant that the antibody has a binding affinity for a non-homologous protein which is at least one order of magnitude, more preferably at least 2 orders of magnitude, and even more preferably at least 3 orders of magnitude less than the binding affinity of the antibody for one or more of the proteins of SEQ ID Nos:8-14.

Following immunization of an animal with an antigenic preparation of a hedgehog protein, anti-*hh* antisera can be obtained and, if desired, polyclonal anti-*hh* antibodies isolated from the serum. To produce monoclonal antibodies, antibody-producing cells (lymphocytes) can be harvested from an immunized animal and fused by standard somatic cell fusion procedures with immortalizing cells such as myeloma cells to yield hybridoma cells. Such techniques are well known in the art, and include, for example, the hybridoma technique (originally developed by Kohler and Milstein, (1975) *Nature*, 256: 495-497), the human B cell hybridoma technique (Kozbar et al., (1983) *Immunology Today*, 4: 72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., (1985) *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. pp. 77-96). Hybridoma cells can be screened immunochemically for production of antibodies specifically reactive with a vertebrate *hh* polypeptide of the present invention and monoclonal antibodies isolated from a culture comprising such hybridoma cells.

The term antibody as used herein is intended to include fragments thereof which are also specifically reactive with one of the subject vertebrate *hh* polypeptides. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. For example, F(ab)₂ fragments can be generated by treating antibody with pepsin. The resulting F(ab)₂ fragment can be treated to reduce disulfide bridges to produce Fab fragments. The antibody of the present invention is further intended to include bispecific and chimeric molecules having affinity for a *hedgehog* protein conferred by at least one CDR region of the antibody.

Both monoclonal and polyclonal antibodies (Ab) directed against authentic *hedgehog* polypeptides, or *hedgehog* variants, and antibody fragments such as Fab and F(ab)₂, can be used to block the action of one or more *hedgehog* proteins and allow the study of the role of these proteins in, for example, embryogenesis and/or maintenance of differential tissue. For example, purified monoclonal Abs can be injected directly into the limb buds of chick or mouse embryos. It is demonstrated in the examples below that *hh* is expressed in the limb buds of, for example, day 10.5 embryos. Thus, the use of anti-*hh* Abs during this developmental stage can allow assessment of the effect of *hh* on the formation of limbs *in vivo*. In a similar approach, hybridomas producing anti-*hh* monoclonal Abs, or biodegradable gels in which anti-*hh* Abs are suspended, can be implanted at a site proximal or within the area at which *hh* action is intended to be blocked. Experiments of this nature

can aid in deciphering the role of this and other factors that may be involved in limb patterning and tissue formation.

Antibodies which specifically bind *hedgehog* epitopes can also be used in immunohistochemical staining of tissue samples in order to evaluate the abundance and pattern of expression of each of the subject *hh* polypeptides. Anti-*hedgehog* antibodies can be used diagnostically in immuno-precipitation and immuno-blotting to detect and evaluate *hedgehog* protein levels in tissue as part of a clinical testing procedure. For instance, such measurements can be useful in predictive valuations of the onset or progression of neurological disorders, such as those marked by denervation-like or disuse-like symptoms. Likewise, the ability to monitor *hh* levels in an individual can allow determination of the efficacy of a given treatment regimen for an individual afflicted with such a disorder. The level of *hh* polypeptides may be measured in bodily fluid, such as in samples of cerebral spinal fluid or amniotic fluid, or can be measured in tissue, such as produced by biopsy. Diagnostic assays using anti-*hh* antibodies can include, for example, immunoassays designed to aid in early diagnosis of a neurodegenerative disorder, particularly ones which are manifest at birth. Diagnostic assays using anti-*hh* polypeptide antibodies can also include immunoassays designed to aid in early diagnosis and phenotyping of a differentiative disorder, as well as neoplastic or hyperplastic disorders.

Another application of anti-*hh* antibodies of the present invention is in the immunological screening of cDNA libraries constructed in expression vectors such as λ gt11, λ gt18-23, λ ZAP, and λ ORF8. Messenger libraries of this type, having coding sequences inserted in the correct reading frame and orientation, can produce fusion proteins. For instance, λ gt11 will produce fusion proteins whose amino termini consist of β -galactosidase amino acid sequences and whose carboxy termini consist of a foreign polypeptide. Antigenic epitopes of an *hh* protein, e.g. other orthologs of a particular *hedgehog* protein or other homologs from the same species, can then be detected with antibodies, as, for example, reacting nitrocellulose filters lifted from infected plates with anti-*hh* antibodies. Positive phage detected by this assay can then be isolated from the infected plate. Thus, the presence of *hedgehog* homologs can be detected and cloned from other animals, as can alternate isoforms (including splicing variants) from humans.

Moreover, the nucleotide sequences determined from the cloning of *hh* genes from vertebrate organisms will further allow for the generation of probes and primers designed for use in identifying and/or cloning *hedgehog* homologs in other cell types, e.g. from other tissues, as well as *hh* homologs from other vertebrate organisms. For instance, the present invention also provides a probe/primer comprising a substantially purified oligonucleotide, which oligonucleotide comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least 10 consecutive nucleotides of sense or anti-sense sequence

selected from the group consisting of SEQ ID No:1, SEQ ID No:2, SEQ ID No:3, SEQ ID No:4, SEQ ID No:5, SEQ ID No:6 and SEQ ID No:7, or naturally occurring mutants thereof. For instance, primers based on the nucleic acid represented in SEQ ID Nos:1-7 can be used in PCR reactions to clone *hedgehog* homologs. Likewise, probes based on the subject *hedgehog* sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto and able to be detected, e.g. the label group is selected from the group consisting of radioisotopes, fluorescent compounds, enzymes, and enzyme co-factors.

Such probes can also be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a *hedgehog* protein, such as by measuring a level of a *hedgehog* encoding nucleic acid in a sample of cells from a patient; e.g. detecting *hh* mRNA levels or determining whether a genomic *hh* gene has been mutated or deleted.

To illustrate, nucleotide probes can be generated from the subject *hedgehog* genes which facilitate histological screening of intact tissue and tissue samples for the presence (or absence) of *hedgehog*-encoding transcripts. Similar to the diagnostic uses of anti-*hedgehog* antibodies, the use of probes directed to *hh* messages, or to genomic *hh* sequences, can be used for both predictive and therapeutic evaluation of allelic mutations which might be manifest in, for example, neoplastic or hyperplastic disorders (e.g. unwanted cell growth) or abnormal differentiation of tissue. Used in conjunction with immunoassays as described above, the oligonucleotide probes can help facilitate the determination of the molecular basis for a developmental disorder which may involve some abnormality associated with expression (or lack thereof) of a *hedgehog* protein. For instance, variation in polypeptide synthesis can be differentiated from a mutation in a coding sequence.

Accordingly, the present method provides a method for determining if a subject is at risk for a disorder characterized by aberrant control of differentiation or unwanted cell proliferation. For instance, the subject assay can be used in the screening and diagnosis of genetic and acquired disorders which involve alteration in one or more of the *hedgehog* genes. In preferred embodiments, the subject method can be generally characterized as comprising: detecting, in a tissue sample of the subject (e.g. a human patient), the presence or absence of a genetic lesion characterized by at least one of (i) a mutation of a gene encoding a *hedgehog* protein or (ii) the mis-expression of a *hedgehog* gene. To illustrate, such genetic lesions can be detected by ascertaining the existence of at least one of (i) a deletion of one or more nucleotides from a *hedgehog* gene, (ii) an addition of one or more nucleotides to a *hedgehog* gene, (iii) a substitution of one or more nucleotides of a *hedgehog* gene, (iv) a gross chromosomal rearrangement of a *hedgehog* gene, (v) a gross alteration in the level of a messenger RNA transcript of an *hh* gene, (vi) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a vertebrate *hh* gene, and (vii) a non-wild

type level of a *hedgehog* protein. In one aspect of the invention there is provided a probe/primer comprising an oligonucleotide containing a region of nucleotide sequence which is capable of hybridizing to a sense or antisense sequence selected from the group consisting of SEQ ID Nos:1-7, or naturally occurring mutants thereof, or 5' or 3' flanking sequences or intronic sequences naturally associated with a vertebrate *hh* gene. The probe is exposed to nucleic acid of a tissue sample; and the hybridization of the probe to the sample nucleic acid is detected. In certain embodiments, detection of the lesion comprises utilizing the probe/primer in a polymerase chain reaction (PCR) (see, e.g., U.S. Patent No: 4,683,195 and 4,683,202) or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran et al. (1988) *Science*, 241:1077-1080; and NaKazawa et al. (1944) *PNAS* 91:360-364) the later of which can be particularly useful for detecting point mutations in *hedgehog* genes. Alternatively, immunoassays can be employed to determine the level of *hh* proteins, either soluble or membrane bound.

Yet another diagnostic screen employs a source of *hedgehog* protein directly. As described herein, *hedgehog* proteins of the present invention are involved in the induction of differentiation. Accordingly, the pathology of certain differentiative and/or proliferative disorders can be marked by loss of *hedgehog* sensitivity by the afflicted tissue. Consequently, the response of a tissue or cell sample to an inductive amount of a *hedgehog* protein can be used to detect and characterize certain cellular transformations and degenerative conditions. For instance, tissue/cell samples from a patient can be treated with a *hedgehog* agonist and the response of the tissue to the treatment determined. Response can be qualified and/or quantified, for example, on the basis of phenotypic change as result of *hedgehog* induction. For example, expression of gene products induced by *hedgehog* treatment can be scored for by immunoassay. The *patched* protein, for example, is upregulated in drosophila in response to Dros-HH, and, in light of the findings herein, a presumed vertebrate homolog will similarly be upregulated. Thus, detection of *patched* expression on the cells of the patient sample can permit detection of tissue that is not *hedgehog*-responsive. Likewise, scoring for other phenotypic markers provides a means for determining the response to *hedgehog*.

Furthermore, by making available purified and recombinant *hedgehog* polypeptides, the present invention facilitates the development of assays which can be used to screen for drugs, including *hedgehog* homologs, which are either agonists or antagonists of the normal cellular function of the subject *hedgehog* polypeptides, or of their role in the pathogenesis of cellular differentiation and/or proliferation and disorders related thereto. In one embodiment, the assay evaluates the ability of a compound to modulate binding between a *hedgehog* polypeptide and a *hedgehog* receptor. A variety of assay formats will suffice and, in light of the present inventions, will be comprehended by skilled artisan.

In many drug screening programs which test libraries of compounds and natural extracts, high throughput assays are desirable in order to maximize the number of compounds surveyed in a given period of time. Assays which are performed in cell-free systems, such as may be derived with purified or semi-purified proteins, are often preferred as "primary" screens in that they can be generated to permit rapid development and relatively easy detection of an alteration in a molecular target which is mediated by a test compound. Moreover, the effects of cellular toxicity and/or bioavailability of the test compound can be generally ignored in the *in vitro* system, the assay instead being focused primarily on the effect of the drug on the molecular target as may be manifest in an alteration of binding affinity with receptor proteins. Accordingly, in an exemplary screening assay of the present invention, the compound of interest is contacted with a *hedgehog* receptor polypeptide which is ordinarily capable of binding a *hedgehog* protein. To the mixture of the compound and receptor is then added a composition containing a *hedgehog* polypeptide. Detection and quantification of receptor/*hedgehog* complexes provides a means for determining the compound's efficacy at inhibiting (or potentiating) complex formation between the receptor protein and the *hedgehog* polypeptide. The efficacy of the compound can be assessed by generating dose response curves from data obtained using various concentrations of the test compound. Moreover, a control assay can also be performed to provide a baseline for comparison. In the control assay, isolated and purified *hedgehog* polypeptide is added to a composition containing the receptor protein, and the formation of receptor/*hedgehog* complex is quantitated in the absence of the test compound.

In an illustrative embodiment, the polypeptide utilized as a *hedgehog* receptor can be generated from the drosophila *patched* protein or a vertebrate homolog thereof. In light of the ability of, for example, *Shh* to activate Dros-HH pathways in transgenic drosophila (see Example 4), it may be concluded that vertebrate *hedgehog* proteins are capable of binding to Drosophila HH receptors. Accordingly, an exemplary screening assay includes a suitable portion of the *patched* protein (SEQ ID No. 42), such as one or both of the substantial extracellular domains (e.g. residues Lys-93 to His-426 and Arg-700 to Arg-966). For instance, the *patched* protein can be provided in soluble form, as for example a preparation of one of the extracellular domains, or a preparation of both of the extracellular domains which are covalently connected by an unstructured linker (see, for example, Huston et al. (1988) PNAS 85:4879; and U.S. Patent No. 5,091,513), or can be provided as part of a liposomal preparation or expressed on the surface of a cell.

Complex formation between the *hedgehog* polypeptide and a *hedgehog* receptor may be detected by a variety of techniques. For instance, modulation of the formation of complexes can be quantitated using, for example, detectably labeled proteins such as

radiolabelled, fluorescently labeled, or enzymatically labeled *hedgehog* polypeptides, by immunoassay, or by chromatographic detection.

Typically, it will be desirable to immobilize either the *hedgehog* receptor or the *hedgehog* polypeptide to facilitate separation of receptor/*hedgehog* complexes from uncomplexed forms of one of the proteins, as well as to accommodate automation of the assay. In one embodiment, a fusion protein can be provided which adds a domain that allows the protein to be bound to a matrix. For example, glutathione-S-transferase/receptor (GST/receptor) fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the *hedgehog* polypeptide, e.g. an ^{35}S -labeled *hedgehog* polypeptide, and the test compound and incubated under conditions conducive to complex formation, e.g. at physiological conditions for salt and pH, though slightly more stringent conditions may be desired. Following incubation, the beads are washed to remove any unbound *hedgehog* polypeptide, and the matrix bead-bound radiolabel determined directly (e.g. beads placed in scintillant), or in the supernatant after the receptor/*hedgehog* complexes are dissociated. Alternatively, the complexes can dissociated from the bead, separated by SDS-PAGE gel, and the level of *hedgehog* polypeptide found in the bead fraction quantitated from the gel using standard electrophoretic techniques.

Other techniques for immobilizing proteins on matrices are also available for use in the subject assay. For instance, soluble portions of the *hedgehog* receptor protein can be immobilized utilizing conjugation of biotin and streptavidin. For instance, biotinylated receptor molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with the *hedgehog* receptor but which do not interfere with *hedgehog* binding can be derivatized to the wells of the plate, and the receptor trapped in the wells by antibody conjugation. As above, preparations of a *hedgehog* polypeptide and a test compound are incubated in the receptor-presenting wells of the plate, and the amount of receptor/*hedgehog* complex trapped in the well can be quantitated. Exemplary methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the *hedgehog* polypeptide, or which are reactive with the receptor protein and compete for binding with the *hedgehog* polypeptide; as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the *hedgehog* polypeptide. In the instance of the latter, the enzyme can be chemically conjugated or provided as a fusion protein with the *hedgehog* polypeptide. To illustrate, the *hedgehog* polypeptide can be chemically cross-linked or genetically fused with alkaline phosphatase, and the amount of *hedgehog*

polypeptide trapped in the complex can be assessed with a chromogenic substrate of the enzyme, e.g. paranitrophenylphosphate. Likewise, a fusion protein comprising the *hedgehog* polypeptide and glutathione-S-transferase can be provided, and complex formation quantitated by detecting the GST activity using 1-chloro-2,4-dinitrobenzene (Habig et al
5 (1974) *J Biol Chem* 249:7130).

For processes which rely on immunodetection for quantitating one of the proteins trapped in the complex, antibodies against the protein, such as the anti-*hedgehog* antibodies described herein, can be used. Alternatively, the protein to be detected in the complex can be "epitope tagged" in the form of a fusion protein which includes, in addition to the *hedgehog*
10 polypeptide or *hedgehog* receptor sequence, a second polypeptide for which antibodies are readily available (e.g. from commercial sources). For instance, the GST fusion proteins described above can also be used for quantification of binding using antibodies against the GST moiety. Other useful epitope tags include myc-epitopes (e.g., see Ellison et al. (1991) *J Biol Chem* 266:21150-21157) which includes a 10-residue sequence from c-myc, as well as
15 the pFLAG system (International Biotechnologies, Inc.) or the pEZZ-protein A system (Pharmacia, NJ).

Where the desired portion of the *hh* receptor (or other *hedgehog* binding molecule) cannot be provided in soluble form, liposomal vesicles can be used to provide manipulatable and isolatable sources of the receptor. For example, both authentic and recombinant forms of
20 the *patched* protein can be reconstituted in artificial lipid vesicles (e.g. phosphatidylcholine liposomes) or in cell membrane-derived vesicles (see, for example, Bear et al. (1992) *Cell* 68:809-818; Newton et al. (1983) *Biochemistry* 22:6110-6117; and Reber et al. (1987) *J Biol Chem* 262:11369-11374).

In addition to cell-free assays, such as described above, the readily available source of
25 vertebrate *hedgehog* proteins provided by the present invention also facilitates the generation of cell-based assays for identifying small molecule agonists/antagonists and the like. Analogous to the cell-based assays described above for screening combinatorial libraries, cells which are sensitive to *hedgehog* induction can be contacted with a *hedgehog* protein and a test agent of interest, with the assay scoring for modulation in *hedgehog* inductive
30 responses by the target cell in the presence and absence of the test agent. As with the cell-free assays, agents which produce a statistically significant change in *hedgehog* activities (either inhibition or potentiation) can be identified. In an illustrative embodiment, motor neuron progenitor cells, such as from neural plate explants, can be used as target cells. Treatment of such explanted cells with, for example, *Shh* causes the cells to differentiate into
35 motor neurons. By detecting the co-expression of the LIM homeodomain protein Islet-1 (Thor et al. (1991) *Neuron* 7:881-889; Ericson et al. (1992) *Science* 256:1555-1560) and the immunoglobulin-like protein SC1 (Tanaka et al. (1984) *Dev Biol* 106:26-37), the ability of a

candidate agent to potentiate or inhibit *Shh* induction of motor neuron differentiation can be measured. The *hedgehog* protein can be provided as a purified source, or in the form of cells/tissue which express the protein and which are co-cultured with the target cells.

In yet another embodiment, the method of the present invention can be used to isolate and clone *hedgehog* receptors. For example, purified *hedgehog* proteins of the present invention can be employed to precipitate *hedgehog* receptor proteins from cell fractions prepared from cells which are responsive to a *hedgehog* protein. For instance, purified *hedgehog* protein can be derivatized with biotin (using, for instance, NHS-Biotin, Pierce Chemical catalog no. 21420G), and the biotinylated protein utilized to saturate membrane bound *hh* receptors. The *hedgehog* bound receptors can subsequently be adsorbed or immobilized on streptavidin. If desired, the *hedgehog*-receptor complex can be cross-linked with a chemical cross-linking agent. In such a manner, *hh* receptors can be purified, preferably to near homogeneity. The isolated *hh* receptor can then be partially digested with, for example, trypsin, and the resulting peptides separated by reverse-phase chromatography. The chromatography fragments are then analyzed by Edman degradation to obtain single sequences for two or more of the proteolytic fragments. From the chemically determined amino acid sequence for each of these tryptic fragments, a set of oligonucleotide primers can be designed for PCR. These primers can be used to screen both genomic and cDNA libraries. Similar strategies for cloning receptors have been employed, for example, to obtain the recombinant gene for somatostatin receptors (Eppler et al. (1992) *J Biol Chem* 267:15603-15612).

Other techniques for identifying *hedgehog* receptors by expression cloning will be evident in light of the present disclosure. For instance, purified *hh* polypeptides can be immobilized in wells of micro titre plates and contacted with, for example, COS cells transfected with a cDNA library (e.g., from tissue expected to be responsive to *hedgehog* induction). From this panning assay, cells which express *hedgehog* receptor molecules can be isolated on the basis of binding to the immobilized *hedgehog* protein. Another cloning system, described in PCT publications WO 92/06220 of Flanagan and Leder, involves the use of an expression cloning system whereby a *hedgehog* receptor is stored on the basis of binding to a *hedgehog*/alkaline phosphatase fusion protein (see also Cheng et al. (1994) *Cell* 79:157-168).

Another aspect of the present invention relates to a method of inducing and/or maintaining a differentiated state, enhancing survival, and/or promoting proliferation of a cell responsive to a vertebrate *hedgehog* protein, by contacting the cells with an *hh* agonist or an *hh* antagonist as the circumstances may warrant. For instance, it is contemplated by the invention that, in light of the present finding of an apparently broad involvement of *hedgehog* proteins in the formation of ordered spatial arrangements of differentiated tissues in

vertebrates, the subject method could be used to generate and/or maintain an array of different vertebrate tissue both *in vitro* and *in vivo*. The *hh* agent, whether inductive or anti-inductive, can be, as appropriate, any of the preparations described above, including isolated polypeptides, gene therapy constructs, antisense molecules, peptidomimetics or agents identified in the drug assays provided herein. Moreover, it is contemplated that, based on the observation of activity of the vertebrate *hedgehog* proteins in *Drosophila*, *hh* agents, for purposes of therapeutic and diagnostic uses, can include the Dros-HH protein and homologs thereof. Moreover, the source of *hedgehog* protein can be, in addition to purified protein or recombinant cells, cells or tissue explants which naturally produce one or more *hedgehog* proteins. For instance, as described in Example 2, neural tube explants from embryos, particularly floorplate tissue, can provide a source for *Shh* polypeptide, which source can be implanted in a patient or otherwise provided, as appropriate, for induction or maintenance of differentiation.

For example, the present method is applicable to cell culture techniques. *In vitro* neuronal culture systems have proved to be fundamental and indispensable tools for the study of neural development, as well as the identification of neurotrophic factors such as nerve growth factor (NGF), ciliary trophic factors (CNTF), and brain derived neurotrophic factor (BDNF). Once a neuronal cell has become terminally-differentiated it typically will not change to another terminally differentiated cell-type. However, neuronal cells can nevertheless readily lose their differentiated state. This is commonly observed when they are grown in culture from adult tissue, and when they form a blastema during regeneration. The present method provides a means for ensuring an adequately restrictive environment in order to maintain neuronal cells at various stages of differentiation, and can be employed, for instance, in cell cultures designed to test the specific activities of other trophic factors. In such embodiments of the subject method, the cultured cells can be contacted with an *hh* polypeptide, or an agent identified in the assays described above, in order to induce neuronal differentiation (e.g. of a stem cell), or to maintain the integrity of a culture of terminally-differentiated neuronal cells by preventing loss of differentiation. The source of *hedgehog* protein in the culture can be derived from, for example, a purified or semi-purified protein composition added directly to the cell culture media, or alternatively, supported and/or released from a polymeric device which supports the growth of various neuronal cells and which has been doped with the protein. The source of the *hedgehog* protein can also be a cell that is co-cultured with the intended neuronal cell and which produces a recombinant wild-type *hedgehog* protein. Alternatively, the source can be the neuronal cell itself which has been engineered to produce a recombinant *hedgehog* protein. In an exemplary embodiment, a naive neuronal cell (e.g. a stem cell) is treated with an *hh* agonist in order to induce differentiation of the cells into, for example, sensory neurons or, alternatively, motoneurons. Such neuronal cultures can be used as convenient assay systems as well as

sources of implantable cells for therapeutic treatments. For example, *hh* polypeptides may be useful in establishing and maintaining the olfactory neuron cultures described in U.S. Patent 5,318,907 and the like.

According to the present invention, large numbers of non-tumorigenic neural progenitor cells can be perpetuated *in vitro* and induced to differentiate by contact with *hedgehog* proteins. Generally, a method is provided comprising the steps of isolating neural progenitor cells from an animal, perpetuating these cells *in vitro* or *in vivo*, preferably in the presence of growth factors, and differentiating these cells into particular neural phenotypes, e.g., neurons and glia, by contacting the cells with a *hedgehog* agonist.

Progenitor cells are thought to be under a tonic inhibitory influence which maintains the progenitors in a suppressed state until their differentiation is required. However, recent techniques have been provided which permit these cells to be proliferated, and unlike neurons which are terminally differentiated and therefore non-dividing, they can be produced in unlimited number and are highly suitable for transplantation into heterologous and autologous hosts with neurodegenerative diseases.

By "progenitor" it is meant an oligopotent or multipotent stem cell which is able to divide without limit and, under specific conditions, can produce daughter cells which terminally differentiate such as into neurons and glia. These cells can be used for transplantation into a heterologous or autologous host. By heterologous is meant a host other than the animal from which the progenitor cells were originally derived. By autologous is meant the identical host from which the cells were originally derived.

Cells can be obtained from embryonic, post-natal, juvenile or adult neural tissue from any animal. By any animal is meant any multicellular animal which contains nervous tissue. More particularly, is meant any fish, reptile, bird, amphibian or mammal and the like. The most preferable donors are mammals, especially mice and humans.

In the case of a heterologous donor animal, the animal may be euthanized, and the brain and specific area of interest removed using a sterile procedure. Brain areas of particular interest include any area from which progenitor cells can be obtained which will serve to restore function to a degenerated area of the host's brain. These regions include areas of the central nervous system (CNS) including the cerebral cortex, cerebellum, midbrain, brainstem, spinal cord and ventricular tissue, and areas of the peripheral nervous system (PNS) including the carotid body and the adrenal medulla. More particularly, these areas include regions in the basal ganglia, preferably the striatum which consists of the caudate and putamen, or various cell groups such as the globus pallidus, the subthalamic nucleus, the nucleus basalis which is found to be degenerated in Alzheimer's Disease patients, or the substantia nigra pars compacta which is found to be degenerated in Parkinson's Disease patients.

53

Human heterologous neural progenitor cells may be derived from fetal tissue obtained from elective abortion, or from a post-natal, juvenile or adult organ donor. Autologous neural tissue can be obtained by biopsy, or from patients undergoing neurosurgery in which neural tissue is removed, in particular during epilepsy surgery, and more particularly during temporal lobectomies and hippocampalectomies.

Cells can be obtained from donor tissue by dissociation of individual cells from the connecting extracellular matrix of the tissue. Dissociation can be obtained using any known procedure, including treatment with enzymes such as trypsin, collagenase and the like, or by using physical methods of dissociation such as with a blunt instrument. Dissociation of fetal cells can be carried out in tissue culture medium, while a preferable medium for dissociation of juvenile and adult cells is artificial cerebral spinal fluid (aCSF). Regular aCSF contains 124 mM NaCl, 5 mM KCl, 1.3 mM MgCl₂, 2 mM CaCl₂, 26 mM NaHCO₃, and 10 mM D-glucose. Low Ca²⁺ aCSF contains the same ingredients except for MgCl₂ at a concentration of 3.2 mM and CaCl₂ at a concentration of 0.1 mM.

Dissociated cells can be placed into any known culture medium capable of supporting cell growth, including MEM, DMEM, RPMI, F-12, and the like, containing supplements which are required for cellular metabolism such as glutamine and other amino acids, vitamins, minerals and useful proteins such as transferrin and the like. Medium may also contain antibiotics to prevent contamination with yeast, bacteria and fungi such as penicillin, streptomycin, gentamicin and the like. In some cases, the medium may contain serum derived from bovine, equine, chicken and the like. A particularly preferable medium for cells is a mixture of DMEM and F-12.

Conditions for culturing should be close to physiological conditions. The pH of the culture media should be close to physiological pH, preferably between pH 6-8, more preferably close to pH 7, even more particularly about pH 7.4. Cells should be cultured at a temperature close to physiological temperature, preferably between 30°C-40°C, more preferably between 32°C-38°C, and most preferably between 35°C-37°C.

Cells can be grown in suspension or on a fixed substrate, but proliferation of the progenitors is preferably done in suspension to generate large numbers of cells by formation of "neurospheres" (see, for example, Reynolds et al. (1992) *Science* 255:1070-1709; and PCT Publications WO93/01275, WO94/09119, WO94/10292, and WO94/16718). In the case of propagating (or splitting) suspension cells, flasks are shaken well and the neurospheres allowed to settle on the bottom corner of the flask. The spheres are then transferred to a 50 ml centrifuge tube and centrifuged at low speed. The medium is aspirated, the cells resuspended in a small amount of medium with growth factor, and the cells mechanically dissociated and resuspended in separate aliquots of media.

54

Cell suspensions in culture medium are supplemented with any growth factor which allows for the proliferation of progenitor cells and seeded in any receptacle capable of sustaining cells, though as set out above, preferably in culture flasks or roller bottles. Cells typically proliferate within 3-4 days in a 37°C incubator, and proliferation can be reinitiated at any time after that by dissociation of the cells and resuspension in fresh medium containing growth factors.

In the absence of substrate, cells lift off the floor of the flask and continue to proliferate in suspension forming a hollow sphere of undifferentiated cells. After approximately 3-10 days *in vitro*, the proliferating clusters (neurospheres) are fed every 2-7 days, and more particularly every 2-4 days by gentle centrifugation and resuspension in medium containing growth factor.

After 6-7 days *in vitro*, individual cells in the neurospheres can be separated by physical dissociation of the neurospheres with a blunt instrument, more particularly by triturating the neurospheres with a pipette. Single cells from the dissociated neurospheres are suspended in culture medium containing growth factors, and differentiation of the cells can be induced by plating (or resuspending) the cells in the presence of a *hedgehog* agonist, and (optionally) any other factor capable of sustaining differentiation, such as bFGF and the like.

To further illustrate other uses of *hedgehog* agonists and antagonists, it is noted that intracerebral grafting has emerged as an additional approach to central nervous system therapies. For example, one approach to repairing damaged brain tissues involves the transplantation of cells from fetal or neonatal animals into the adult brain (Dunnett et al. (1987) *J Exp Biol* 123:265-289; and Freund et al. (1985) *J Neurosci* 5:603-616). Fetal neurons from a variety of brain regions can be successfully incorporated into the adult brain, and such grafts can alleviate behavioral defects. For example, movement disorder induced by lesions of dopaminergic projections to the basal ganglia can be prevented by grafts of embryonic dopaminergic neurons. Complex cognitive functions that are impaired after lesions of the neocortex can also be partially restored by grafts of embryonic cortical cells. The use of *hedgehog* proteins or mimetics, such as *Shh* or *Dhh*, in the culture can prevent loss of differentiation, or where fetal tissue is used, especially neuronal stem cells, can be used to induce differentiation.

Stem cells useful in the present invention are generally known. For example, several neural crest cells have been identified, some of which are multipotent and likely represent uncommitted neural crest cells, and others of which can generate only one type of cell, such as sensory neurons, and likely represent committed progenitor cells. The role of *hedgehog* proteins employed in the present method to culture such stem cells can be to induce differentiation of the uncommitted progenitor and thereby give rise to a committed progenitor

55

cell, or to cause further restriction of the developmental fate of a committed progenitor cell towards becoming a terminally-differentiated neuronal cell. For example, the present method can be used *in vitro* to induce and/or maintain the differentiation of neural crest cells into glial cells, schwann cells, chromaffin cells, cholinergic sympathetic or parasympathetic neurons, as well as peptidergic and serotonergic neurons. The *hedgehog* protein can be used alone, or can be used in combination with other neurotrophic factors which act to more particularly enhance a particular differentiation fate of the neuronal progenitor cell. In the later instance, an *hh* polypeptide might be viewed as ensuring that the treated cell has achieved a particular phenotypic state such that the cell is poised along a certain developmental pathway so as to be properly induced upon contact with a secondary neurotrophic factor. In similar fashion, even relatively undifferentiated stem cells or primitive neuroblasts can be maintained in culture and caused to differentiate by treatment with *hedgehog* agonists. Exemplary primitive cell cultures comprise cells harvested from the neural plate or neural tube of an embryo even before much overt differentiation has occurred.

In addition to the implantation of cells cultured in the presence of a functional *hedgehog* activity and other *in vitro* uses described above, yet another aspect of the present invention concerns the therapeutic application of a *hedgehog* protein or mimetic to enhance survival of neurons and other neuronal cells in both the central nervous system and the peripheral nervous system. The ability of *hedgehog* protein to regulate neuronal differentiation during development of the nervous system and also presumably in the adult state indicates that certain of the *hedgehog* proteins can be reasonably expected to facilitate control of adult neurons with regard to maintenance, functional performance, and aging of normal cells; repair and regeneration processes in chemically or mechanically lesioned cells; and prevention of degeneration and premature death which result from loss of differentiation in certain pathological conditions. In light of this understanding, the present invention specifically contemplates applications of the subject method to the treatment of (prevention and/or reduction of the severity of) neurological conditions deriving from: (i) acute, subacute, or chronic injury to the nervous system, including traumatic injury, chemical injury, vasal injury and deficits (such as the ischemia resulting from stroke), together with infectious/inflammatory and tumor-induced injury; (ii) aging of the nervous system including Alzheimer's disease; (iii) chronic neurodegenerative diseases of the nervous system, including Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis and the like, as well as spinocerebellar degenerations; and (iv) chronic immunological diseases of the nervous system or affecting the nervous system, including multiple sclerosis.

Many neurological disorders are associated with degeneration of discrete populations of neuronal elements and may be treatable with a therapeutic regimen which includes a *hedgehog* agonist. For example, Alzheimer's disease is associated with deficits in several

neurotransmitter systems, both those that project to the neocortex and those that reside with the cortex. For instance, the nucleus basalis in patients with Alzheimer's disease have been observed to have a profound (75%) loss of neurons compared to age-matched controls. Although Alzheimer's disease is by far the most common form of dementia, several other disorders can produce dementia. Several of these are degenerative diseases characterized by the death of neurons in various parts of the central nervous system, especially the cerebral cortex. However, some forms of dementia are associated with degeneration of the thalamus or the white matter underlying the cerebral cortex. Here, the cognitive dysfunction results from the isolation of cortical areas by the degeneration of efferents and afferents. Huntington's disease involves the degeneration of intrastriatal and cortical cholinergic neurons and GABAergic neurons. Pick's disease is a severe neuronal degeneration in the neocortex of the frontal and anterior temporal lobes, sometimes accompanied by death of neurons in the striatum. Treatment of patients suffering from such degenerative conditions can include the application of *hedgehog* polypeptides, or agents which mimic their effects, in order to control, for example, differentiation and apoptotic events which give rise to loss of neurons (e.g. to enhance survival of existing neurons) as well as promote differentiation and repopulation by progenitor cells in the area affected. In preferred embodiments, a source of a *hedgehog* agent is stereotactically provided within or proximate the area of degeneration. In addition to degenerative-induced dementias, a pharmaceutical preparation of one or more of the subject *hedgehog* proteins can be applied opportunely in the treatment of neurodegenerative disorders which have manifestations of tremors and involuntary movements. Parkinson's disease, for example, primarily affects subcortical structures and is characterized by degeneration of the nigrostriatal pathway, raphe nuclei, locus cereleus, and the motor nucleus of vagus. Ballism is typically associated with damage to the subthalamic nucleus, often due to acute vascular accident. Also included are neurogenic and myopathic diseases which ultimately affect the somatic division of the peripheral nervous system and are manifest as neuromuscular disorders. Examples include chronic atrophies such as amyotrophic lateral sclerosis, Guillain-Barre syndrome and chronic peripheral neuropathy, as well as other diseases which can be manifest as progressive bulbar palsies or spinal muscular atrophies. The present method is amenable to the treatment of disorders of the cerebellum which result in hypotonia or ataxia, such as those lesions in the cerebellum which produce disorders in the limbs ipsilateral to the lesion. For instance, a preparation of a *hedgehog* homolog can be used to treat a restricted form of cerebellar cortical degeneration involving the anterior lobes (vermis and leg areas) such as is common in alcoholic patients.

In an illustrative embodiment, the subject method is used to treat amyotrophic lateral sclerosis. ALS is a name given to a complex of disorders that comprise upper and lower motor neurons. Patients may present with progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, or a combination of these conditions. The major

pathological abnormality is characterized by a selective and progressive degeneration of the lower motor neurons in the spinal cord and the upper motor neurons in the cerebral cortex. The therapeutic application of a *hedgehog* agonist, particularly *Dhh*, can be used alone, or in conjunction with other neurotrophic factors such as CNTF, BDNF or NGF to prevent and/or reverse motor neuron degeneration in ALS patients.

Hedgehog proteins of the present invention can also be used in the treatment of autonomic disorders of the peripheral nervous system, which include disorders affecting the innervation of smooth muscle and endocrine tissue (such as glandular tissue). For instance, the subject method can be used to treat tachycardia or atrial cardiac arrhythmias which may arise from a degenerative condition of the nerves innervating the striated muscle of the heart.

Furthermore, a potential role for certain of the *hedgehog* proteins, which is apparent from the appended examples, mainly the data of respecting *hedgehog* expression in sensory and motor neurons of the head and trunk (including limb buds), concerns the role of *hedgehog* proteins in development and maintenance of dendritic processes of axonal neurons. Potential roles for *hedgehog* proteins consequently include guidance for axonal projections and the ability to promote differentiation and/or maintenance of the innervating cells to their axonal processes. Accordingly, compositions comprising *hedgehog* agonists or other *hedgehog* agents described herein, may be employed to support, or alternatively antagonize the survival and reprojection of several types of ganglionic neurons sympathetic and sensory neurons as well as motor neurons. In particular, such therapeutic compositions may be useful in treatments designed to rescue, for example, various neurons from lesion-induced death as well as guiding reprojection of these neurons after such damage. Such diseases include, but are not limited to, CNS trauma infarction, infection (such as viral infection with varicella-zoster), metabolic disease, nutritional deficiency, toxic agents (such as cisplatin treatment). Moreover, certain of the *hedgehog* agents (such as antagonistic form) may be useful in the selective ablation of sensory neurons, for example, in the treatment of chronic pain syndromes.

As appropriate, *hedgehog* agents can be used in nerve prostheses for the repair of central and peripheral nerve damage. In particular, where a crushed or severed axon is intubulated by use of a prosthetic device, *hedgehog* polypeptides can be added to the prosthetic device to increase the rate of growth and regeneration of the dendritic processes. Exemplary nerve guidance channels are described in U.S. patents 5,092,871 and 4,955,892. Accordingly, a severed axonal process can be directed toward the nerve ending from which it was severed by a prosthesis nerve guide which contains, e.g. a semi-solid formulation containing *hedgehog* polypeptide or mimetic, or which is derivatized along the inner walls with a *hedgehog* protein.

58

In another embodiment, the subject method can be used in the treatment of neoplastic or hyperplastic transformations such as may occur in the central nervous system. For instance, certain of the *hedgehog* proteins (or *hh* agonists) which induce differentiation of neuronal cells can be utilized to cause such transformed cells to become either post-mitotic or apoptotic. Treatment with a *hedgehog* agent may facilitate disruption of autocrine loops, such as TGF- β or PDGF autostimulatory loops, which are believed to be involved in the neoplastic transformation of several neuronal tumors. *Hedgehog* agonists may, therefore, be of use in the treatment of, for example, malignant gliomas, medulloblastomas, neuroectodermal tumors, and ependymomas.

Yet another aspect of the present invention concerns the application of the discovery that *hedgehog* proteins are morphogenic signals involved in other vertebrate organogenic pathways in addition to neuronal differentiation as described above, having apparent roles in other endodermal patterning, as well as both mesodermal and endodermal differentiation processes. As described in the Examples below, *Shh* clearly plays a role in proper limb growth and patterning by initiating expression of signaling molecules, including *Bmp-2* in the mesoderm and *Fgf-4* in the ectoderm. Thus, it is contemplated by the invention that compositions comprising *hedgehog* proteins can also be utilized for both cell culture and therapeutic methods involving generation and maintenance of non-neuronal tissue.

In one embodiment, the present invention makes use of the discovery that *hedgehog* proteins, such as *Shh*, are apparently involved in controlling the development of stem cells responsible for formation of the digestive tract, liver, lungs, and other organs which derive from the primitive gut. As described in the Examples below, *Shh* serves as an inductive signal from the endoderm to the mesoderm, which is critical to gut morphogenesis. Therefore, for example, *hedgehog* agonists can be employed in the development and maintenance of an artificial liver which can have multiple metabolic functions of a normal liver. In an exemplary embodiment, *hedgehog* agonists can be used to induce differentiation of digestive tube stem cells to form hepatocyte cultures which can be used to populate extracellular matrices, or which can be encapsulated in biocompatible polymers, to form both implantable and extracorporeal artificial livers.

In another embodiment, therapeutic compositions of *hedgehog* agonists can be utilized in conjunction with transplantation of such artificial livers, as well as embryonic liver structures, to promote intraperitoneal implantation, vascularization, and *in vivo* differentiation and maintenance of the engrafted liver tissue.

In yet another embodiment, *hedgehog* agonists can be employed therapeutically to regulate such organs after physical, chemical or pathological insult. For instance, therapeutic compositions comprising *hedgehog* agonists can be utilized in liver repair subsequent to a

59

partial hepatectomy. Similarly, therapeutic compositions containing *hedgehog* agonists can be used to promote regeneration of lung tissue in the treatment of emphysema.

In still another embodiment of the present invention, compositions comprising *hedgehog* agonists can be used in the *in vitro* generation of skeletal tissue, such as from
5 skeletogenic stem cells, as well as the *in vivo* treatment of skeletal tissue deficiencies. The present invention particularly contemplates the use of *hedgehog* agonists which maintain a skeletogenic activity, such as an ability to induce chondrogenesis and/or osteogenesis. By "skeletal tissue deficiency", it is meant a deficiency in bone or other skeletal connective tissue at any site where it is desired to restore the bone or connective tissue, no matter how the
10 deficiency originated, e.g. whether as a result of surgical intervention, removal of tumor, ulceration, implant, fracture, or other traumatic or degenerative conditions.

For instance, the present invention makes available effective therapeutic methods and compositions for restoring cartilage function to a connective tissue. Such methods are useful in, for example, the repair of defects or lesions in cartilage tissue which is the result of
15 degenerative wear such as that which results in arthritis, as well as other mechanical derangements which may be caused by trauma to the tissue, such as a displacement of torn meniscus tissue, meniscectomy, a laxation of a joint by a torn ligament, malignment of joints, bone fracture, or by hereditary disease. The present reparative method is also useful for remodeling cartilage matrix, such as in plastic or reconstructive surgery, as well as
20 periodontal surgery. The present method may also be applied to improving a previous reparative procedure, for example, following surgical repair of a meniscus, ligament, or cartilage. Furthermore, it may prevent the onset or exacerbation of degenerative disease if applied early enough after trauma.

In one embodiment of the present invention, the subject method comprises treating
25 the afflicted connective tissue with a therapeutically sufficient amount of a *hedgehog* agonist, particularly an *Ihh* agonist, to generate a cartilage repair response in the connective tissue by stimulating the differentiation and/or proliferation of chondrocytes embedded in the tissue. Induction of chondrocytes by treatment with a *hedgehog* agonist can subsequently result in the synthesis of new cartilage matrix by the treated cells. Such connective tissues as articular
30 cartilage, interarticular cartilage (menisci), costal cartilage (connecting the true ribs and the sternum), ligaments, and tendons are particularly amenable to treatment in reconstructive and/or regenerative therapies using the subject method. As used herein, regenerative therapies include treatment of degenerative states which have progressed to the point of which impairment of the tissue is obviously manifest, as well as preventive treatments of
35 tissue where degeneration is in its earliest stages or imminent. The subject method can further be used to prevent the spread of mineralisation into fibrotic tissue by maintaining a constant production of new cartilage.

60

In an illustrative embodiment, the subject method can be used to treat cartilage of a diarthroidal joint, such as a knee, an ankle, an elbow, a hip, a wrist, a knuckle of either a finger or toe, or a temporomandibular joint. The treatment can be directed to the meniscus of the joint, to the articular cartilage of the joint, or both. To further illustrate, the subject method can be used to treat a degenerative disorder of a knee, such as which might be the result of traumatic injury (e.g., a sports injury or excessive wear) or osteoarthritis. An injection of a *hedgehog* agonist into the joint with, for instance, an arthroscopic needle, can be used to treat the afflicted cartilage. In some instances, the injected agent can be in the form of a hydrogel or other slow release vehicle described above in order to permit a more extended and regular contact of the agent with the treated tissue.

The present invention further contemplates the use of the subject method in the field of cartilage transplantation and prosthetic device therapies. To date, the growth of new cartilage from either transplantation of autologous or allogenic cartilage has been largely unsuccessful. Problems arise, for instance, because the characteristics of cartilage and fibrocartilage varies between different tissue: such as between articular, meniscal cartilage, ligaments, and tendons, between the two ends of the same ligament or tendon, and between the superficial and deep parts of the tissue. The zonal arrangement of these tissues may reflect a gradual change in mechanical properties, and failure occurs when implanted tissue, which has not differentiated under those conditions, lacks the ability to appropriately respond. For instance, when meniscal cartilage is used to repair anterior cruciate ligaments, the tissue undergoes a metaplasia to pure fibrous tissue. By promoting chondrogenesis, the subject method can be used to particularly addresses this problem, by causing the implanted cells to become more adaptive to the new environment and effectively resemble hypertrophic chondrocytes of an earlier developmental stage of the tissue. Thus, the action of chondrogenesis in the implanted tissue, as provided by the subject method, and the mechanical forces on the actively remodeling tissue can synergize to produce an improved implant more suitable for the new function to which it is to be put.

In similar fashion, the subject method can be applied to enhancing both the generation of prosthetic cartilage devices and to their implantation. The need for improved treatment has motivated research aimed at creating new cartilage that is based on collagen-glycosaminoglycan templates (Stone et al. (1990) *Clin Orthop Relat Res* 252:129), isolated chondrocytes (Grande et al. (1989) *J Orthop Res* 7:208; and Takigawa et al. (1987) *Bone Miner* 2:449), and chondrocytes attached to natural or synthetic polymers (Walitani et al. (1989) *J Bone Jt Surg* 71B:74; Vacanti et al. (1991) *Plast Reconstr Surg* 88:753; von Schroeder et al. (1991) *J Biomed Mater Res* 25:329; Freed et al. (1993) *J Biomed Mater Res* 27:11; and the Vacanti et al. U.S. Patent No. 5,041,138). For example, chondrocytes can be grown in culture on biodegradable, biocompatible highly porous scaffolds formed from

polymers such as polyglycolic acid, polylactic acid, agarose gel, or other polymers which degrade over time as function of hydrolysis of the polymer backbone into innocuous monomers. The matrices are designed to allow adequate nutrient and gas exchange to the cells until engraftment occurs. The cells can be cultured *in vitro* until adequate cell volume and density has developed for the cells to be implanted. One advantage of the matrices is that they can be cast or molded into a desired shape on an individual basis, so that the final product closely resembles the patient's own ear or nose (by way of example), or flexible matrices can be used which allow for manipulation at the time of implantation, as in a joint.

In one embodiment of the subject method, the implants are contacted with a *hedgehog* agonist during the culturing process, such as an *Ihh* agonist, in order to induce and/or maintain differentiated chondrocytes in the culture in order as to further stimulate cartilage matrix production within the implant. In such a manner, the cultured cells can be caused to maintain a phenotype typical of a chondrogenic cell (i.e. hypertrophic), and hence continue the population of the matrix and production of cartilage tissue.

In another embodiment, the implanted device is treated with a *hedgehog* agonist in order to actively remodel the implanted matrix and to make it more suitable for its intended function. As set out above with respect to tissue transplants, the artificial transplants suffer from the same deficiency of not being derived in a setting which is comparable to the actual mechanical environment in which the matrix is implanted. The activation of the chondrocytes in the matrix by the subject method can allow the implant to acquire characteristics similar to the tissue for which it is intended to replace.

In yet another embodiment, the subject method is used to enhance attachment of prosthetic devices. To illustrate, the subject method can be used in the implantation of a periodontal prosthesis, wherein the treatment of the surrounding connective tissue stimulates formation of periodontal ligament about the prosthesis, as well as inhibits formation of fibrotic tissue proximate the prosthetic device.

In still further embodiments, the subject method can be employed for the generation of bone (osteogenesis) at a site in the animal where such skeletal tissue is deficient. Indian *hedgehog* is particularly associated with the hypertrophic chondrocytes that are ultimately replaced by osteoblasts. For instance, administration of a *hedgehog* agent of the present invention can be employed as part of a method for treating bone loss in a subject, e.g. to prevent and/or reverse osteoporosis and other osteopenic disorders, as well as to regulate bone growth and maturation. For example, preparations comprising *hedgehog* agonists can be employed, for example, to induce endochondral ossification, at least so far as to facilitate the formation of cartilaginous tissue precursors to form the "model" for ossification. Therapeutic compositions of *hedgehog* agonists can be supplemented, if required, with other

osteoinductive factors, such as bone growth factors (e.g. TGF- β factors, such as the bone morphogenetic factors *BMP-2* and *BMP-4*, as well as activin), and may also include, or be administered in combination with, an inhibitor of bone resorption such as estrogen, bisphosphonate, sodium fluoride, calcitonin, or tamoxifen, or related compounds. However, it will be appreciated that *hedgehog* proteins, such as *Ihh* and *Shh* are likely to be upstream of BMPs, e.g. *hh* treatment will have the advantage of initiating endogenous expression of BMPs along with other factors.

In yet another embodiment of the present invention, a *hedgehog* antagonist can be used to inhibit spermatogenesis. Thus, in light of the present finding that *hedgehog* proteins are involved in the differentiation and/or proliferation and maintenance of testicular germ cells, *hedgehog* antagonist can be utilized to block the action of a naturally-occurring *hedgehog* protein. In a preferred embodiment, the *hedgehog* antagonist inhibits the biological activity of *Dhh* with respect to spermatogenesis, by competitively binding *hedgehog* receptors in the testis. In similar fashion, *hedgehog* agonists and antagonists are potentially useful for modulating normal ovarian function.

The source of *hedgehog* polypeptides, whether for cell culture or for *in vivo* application, can be in the form of a purified protein composition, or can be from a cell expressing either a recombinant or endogenous form of the polypeptide, such as embryonic tissue (e.g., floor plate tissue explants). Moreover, in addition to those forms of the vertebrate *hedgehog* polypeptides described herein, the present invention further contemplates the use of the *drosophila* *hedgehog* (Dros-HH) protein to induce cells and tissue of vertebrate organisms.

In the instance of protein compositions, the *hedgehog* protein, or a pharmaceutically acceptable salt thereof, may be conveniently formulated for administration with a biologically acceptable medium, such as water, buffered saline, polyol (for example, glycerol, propylene glycol, liquid polyethylene glycol and the like) or suitable mixtures thereof. The optimum concentration of the active ingredient(s) in the chosen medium can be determined empirically, according to procedures well known to medicinal chemists. As used herein, "biologically acceptable medium" includes any and all solvents, dispersion media, and the like which may be appropriate for the desired route of administration of the pharmaceutical preparation. The use of such media for pharmaceutically active substances is known in the art. Except insofar as any conventional media or agent is incompatible with the activity of the *hedgehog* protein, its use in the pharmaceutical preparation of the invention is contemplated. Suitable vehicles and their formulation inclusive of other proteins are described, for example, in the book *Remington's Pharmaceutical Sciences* (Remington's Pharmaceutical Sciences. Mack Publishing Company, Easton, Pa., USA 1985). These vehicles include injectable "deposit formulations". Based on the above, such pharmaceutical

formulations include, although not exclusively, solutions or freeze-dried powders of a *hedgehog* homolog (such as a *Shh*, *Dhh* or *Mhh*) in association with one or more pharmaceutically acceptable vehicles or diluents, and contained in buffered media at a suitable pH and isosmotic with physiological fluids. For illustrative purposes only and without being limited by the same, possible compositions or formulations which may be prepared in the form of solutions for the treatment of nervous system disorders with a *hedgehog* protein are given in U.S. Patent No. 5,218,094. In the case of freeze-dried preparations, supporting excipients such as, but not exclusively, mannitol or glycine may be used and appropriate buffered solutions of the desired volume will be provided so as to obtain adequate isotonic buffered solutions of the desired pH. Similar solutions may also be used for the pharmaceutical compositions of *hh* in isotonic solutions of the desired volume and include, but not exclusively, the use of buffered saline solutions with phosphate or citrate at suitable concentrations so as to obtain at all times isotonic pharmaceutical preparations of the desired pH, (for example, neutral pH).

Methods of introduction of exogenous *hh* at the site of treatment include, but are not limited to, intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, oral, intranasal and topical. In addition, it may be desirable to introduce the pharmaceutical compositions of the invention into the central nervous system by any suitable route, including intraventricular and intrathecal injection. Intraventricular injection may be facilitated by an intraventricular catheter, for example, attached to a reservoir, such as an Ommaya reservoir.

Methods of introduction may also be provided by rechargeable or biodegradable devices. Various slow release polymeric devices have been developed and tested *in vivo* in recent years for the controlled delivery of drugs, including proteinacious biopharmaceuticals. A variety of biocompatible polymers (including hydrogels), including both biodegradable and non-degradable polymers, can be used to form an implant for the sustained release of an *hh* at a particular target site. Such embodiments of the present invention can be used for the delivery of an exogenously purified *hedgehog* protein, which has been incorporated in the polymeric device, or for the delivery of *hedgehog* produced by a cell encapsulated in the polymeric device.

An essential feature of certain embodiments of the implant can be the linear release of the *hh*, which can be achieved through the manipulation of the polymer composition and form. By choice of monomer composition or polymerization technique, the amount of water, porosity and consequent permeability characteristics can be controlled. The selection of the shape, size, polymer, and method for implantation can be determined on an individual basis according to the disorder to be treated and the individual patient response. The generation of such implants is generally known in the art. See, for example, *Concise Encyclopedia of Medical & Dental Materials*, ed. by David Williams (MIT Press: Cambridge, MA, 1990);

and the Sabel et al. U.S. Patent No. 4,883,666. In another embodiment of an implant, a source of cells producing a *hedgehog* protein, or a solution of hydrogel matrix containing purified *hh*, is encapsulated in implantable hollow fibers. Such fibers can be pre-spun and subsequently loaded with the *hedgehog* source (Aebischer et al. U.S. Patent No. 4,892,538; Aebischer et al. U.S. Patent No. 5,106,627; Hoffman et al. (1990) *Expt. Neurobiol.* 110:39-44; Jaeger et al. (1990) *Prog. Brain Res.* 82:41-46; and Aebischer et al. (1991) *J. Biomech. Eng.* 113:178-183), or can be co-extruded with a polymer which acts to form a polymeric coat about the *hh* source (Lim U.S. Patent No. 4,391,909; Sefton U.S. Patent No. 4,353,888; Sugamori et al. (1989) *Trans. Am. Artif. Intern. Organs* 35:791-799; Sefton et al. (1987) *Biotechnol. Bioeng.* 29:1135-1143; and Aebischer et al. (1991) *Biomaterials* 12:50-55).

In yet another embodiment of the present invention, the pharmaceutical *hedgehog* protein can be administered as part of a combinatorial therapy with other agents. For example, the combinatorial therapy can include a *hedgehog* protein with at least one trophic factor. Exemplary trophic factors include nerve growth factor, ciliary neurotrophic growth factor, schwannoma-derived growth factor, glial growth factor, stiatal-derived neuronotrophic factor, platelet-derived growth factor, and scatter factor (HGF-SF). Antimitogenic agents can also be used, for example, when proliferation of surrounding glial cells or astrocytes is undesirable in the regeneration of nerve cells. Examples of such antimitotic agents include cytosine, arabinoside, 5-fluorouracil, hydroxyurea, and methotrexate.

Another aspect of the invention features transgenic non-human animals which express a heterologous *hedgehog* gene of the present invention, or which have had one or more genomic *hedgehog* genes disrupted in at least one of the tissue or cell-types of the animal. Accordingly, the invention features an animal model for developmental diseases, which animal has *hedgehog* allele which is mis-expressed. For example, a mouse can be bred which has one or more *hh* alleles deleted or otherwise rendered inactive. Such a mouse model can then be used to study disorders arising from mis-expressed *hedgehog* genes, as well as for evaluating potential therapies for similar disorders.

Another aspect of the present invention concerns transgenic animals which are comprised of cells (of that animal) which contain a transgene of the present invention and which preferably (though optionally) express an exogenous *hedgehog* protein in one or more cells in the animal. A *hedgehog* transgene can encode the wild-type form of the protein, or can encode homologs thereof, including both agonists and antagonists, as well as antisense constructs. In preferred embodiments, the expression of the transgene is restricted to specific subsets of cells, tissues or developmental stages utilizing, for example, cis-acting sequences that control expression in the desired pattern. In the present invention, such mosaic expression of a *hedgehog* protein can be essential for many forms of lineage analysis and can additionally provide a means to assess the effects of, for example, lack of *hedgehog*

65

expression which might grossly alter development in small patches of tissue within an otherwise normal embryo. Toward this end, tissue-specific regulatory sequences and conditional regulatory sequences can be used to control expression of the transgene in certain spatial patterns. Moreover, temporal patterns of expression can be provided by, for example, conditional recombination systems or prokaryotic transcriptional regulatory sequences.

Genetic techniques which allow for the expression of transgenes can be regulated via site-specific genetic manipulation *in vivo* are known to those skilled in the art. For instance, genetic systems are available which allow for the regulated expression of a recombinase that catalyzes the genetic recombination a target sequence. As used herein, the phrase "target sequence" refers to a nucleotide sequence that is genetically recombined by a recombinase. The target sequence is flanked by recombinase recognition sequences and is generally either excised or inverted in cells expressing recombinase activity. Recombinase catalyzed recombination events can be designed such that recombination of the target sequence results in either the activation or repression of expression of one of the subject *hedgehog* proteins. For example, excision of a target sequence which interferes with the expression of a recombinant *hh* gene, such as one which encodes an antagonistic homolog or an antisense transcript, can be designed to activate expression of that gene. This interference with expression of the protein can result from a variety of mechanisms, such as spatial separation of the *hh* gene from the promoter element or an internal stop codon. Moreover, the transgene can be made wherein the coding sequence of the gene is flanked by recombinase recognition sequences and is initially transfected into cells in a 3' to 5' orientation with respect to the promoter element. In such an instance, inversion of the target sequence will reorient the subject gene by placing the 5' end of the coding sequence in an orientation with respect to the promoter element which allow for promoter driven transcriptional activation.

In an illustrative embodiment, either the *cre/loxP* recombinase system of bacteriophage P1 (Lakso et al. (1992) *PNAS* 89:6232-6236; Orban et al. (1992) *PNAS* 89:6861-6865) or the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman et al. (1991) *Science* 251:1351-1355; PCT publication WO 92/15694) can be used to generate *in vivo* site-specific genetic recombination systems. Cre recombinase catalyzes the site-specific recombination of an intervening target sequence located between *loxP* sequences. *loxP* sequences are 34 base pair nucleotide repeat sequences to which the Cre recombinase binds and are required for Cre recombinase mediated genetic recombination. The orientation of *loxP* sequences determines whether the intervening target sequence is excised or inverted when Cre recombinase is present (Abremski et al. (1984) *J. Biol. Chem.* 259:1509-1514); catalyzing the excision of the target sequence when the *loxP* sequences are oriented as direct repeats and catalyzes inversion of the target sequence when *loxP* sequences are oriented as inverted repeats.

66

Accordingly, genetic recombination of the target sequence is dependent on expression of the Cre recombinase. Expression of the recombinase can be regulated by promoter elements which are subject to regulatory control, e.g., tissue-specific, developmental stage-specific, inducible or repressible by externally added agents. This regulated control will result in genetic recombination of the target sequence only in cells where recombinase expression is mediated by the promoter element. Thus, the activation expression of a recombinant *hedgehog* protein can be regulated via control of recombinase expression.

Use of the *cre/loxP* recombinase system to regulate expression of a recombinant *hh* protein requires the construction of a transgenic animal containing transgenes encoding both the Cre recombinase and the subject protein. Animals containing both the Cre recombinase and a recombinant *hedgehog* gene can be provided through the construction of "double" transgenic animals. A convenient method for providing such animals is to mate two transgenic animals each containing a transgene, e.g., an *hh* gene and recombinase gene.

One advantage derived from initially constructing transgenic animals containing a *hedgehog* transgene in a recombinase-mediated expressible format derives from the likelihood that the subject protein, whether agonistic or antagonistic, can be deleterious upon expression in the transgenic animal. In such an instance, a founder population, in which the subject transgene is silent in all tissues, can be propagated and maintained. Individuals of this founder population can be crossed with animals expressing the recombinase in, for example, one or more tissues and/or a desired temporal pattern. Thus, the creation of a founder population in which, for example, an antagonistic *hh* transgene is silent will allow the study of progeny from that founder in which disruption of *hedgehog* mediated induction in a particular tissue or at certain developmental stages would result in, for example, a lethal phenotype.

Similar conditional transgenes can be provided using prokaryotic promoter sequences which require prokaryotic proteins to be simultaneously expressed in order to facilitate expression of the *hedgehog* transgene. Exemplary promoters and the corresponding trans-activating prokaryotic proteins are given in U.S. Patent No. 4,833,080.

Moreover, expression of the conditional transgenes can be induced by gene therapy-like methods wherein a gene encoding the trans-activating protein, e.g. a recombinase or a prokaryotic protein, is delivered to the tissue and caused to be expressed, such as in a cell-type specific manner. By this method, a *hedgehog* transgene could remain silent into adulthood until "turned on" by the introduction of the trans-activator.

In an exemplary embodiment, the "transgenic non-human animals" of the invention are produced by introducing transgenes into the germline of the non-human animal. Embryonic target cells at various developmental stages can be used to introduce transgenes.

Different methods are used depending on the stage of development of the embryonic target cell. The zygote is the best target for micro-injection. In the mouse, the male pronucleus reaches the size of approximately 20 micrometers in diameter which allows reproducible injection of 1-2pl of DNA solution. The use of zygotes as a target for gene transfer has a major advantage in that in most cases the injected DNA will be incorporated into the host gene before the first cleavage (Brinster et al. (1985) *PNAS* 82:4438-4442). As a consequence, all cells of the transgenic non-human animal will carry the incorporated transgene. This will in general also be reflected in the efficient transmission of the transgene to offspring of the founder since 50% of the germ cells will harbor the transgene. Microinjection of zygotes is the preferred method for incorporating transgenes in practicing the invention.

Retroviral infection can also be used to introduce *hedgehog* transgenes into a non-human animal. The developing non-human embryo can be cultured *in vitro* to the blastocyst stage. During this time, the blastomeres can be targets for retroviral infection (Jaenich, R. (1976) *PNAS* 73:1260-1264). Efficient infection of the blastomeres is obtained by enzymatic treatment to remove the zona pellucida (*Manipulating the Mouse Embryo*, Hogan eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1986). The viral vector system used to introduce the transgene is typically a replication-defective retrovirus carrying the transgene (Jahner et al. (1985) *PNAS* 82:6927-6931; Van der Putten et al. (1985) *PNAS* 82:6148-6152). Transfection is easily and efficiently obtained by culturing the blastomeres on a monolayer of virus-producing cells (Van der Putten, *supra*; Stewart et al. (1987) *EMBO J.* 6:383-388). Alternatively, infection can be performed at a later stage. Virus or virus-producing cells can be injected into the blastocoele (Jahner et al. (1982) *Nature* 298:623-628). Most of the founders will be mosaic for the transgene since incorporation occurs only in a subset of the cells which formed the transgenic non-human animal. Further, the founder may contain various retroviral insertions of the transgene at different positions in the genome which generally will segregate in the offspring. In addition, it is also possible to introduce transgenes into the germ line by intrauterine retroviral infection of the midgestation embryo (Jahner et al. (1982) *supra*).

A third type of target cell for transgene introduction is the embryonic stem cell (ES). ES cells are obtained from pre-implantation embryos cultured *in vitro* and fused with embryos (Evans et al. (1981) *Nature* 292:154-156; Bradley et al. (1984) *Nature* 309:255-258; Gossler et al. (1986) *PNAS* 83: 9065-9069; and Robertson et al. (1986) *Nature* 322:445-448). Transgenes can be efficiently introduced into the ES cells by DNA transfection or by retrovirus-mediated transduction. Such transformed ES cells can thereafter be combined with blastocysts from a non-human animal. The ES cells thereafter colonize the embryo and contribute to the germ line of the resulting chimeric animal. For review see Jaenisch, R. (1988) *Science* 240:1468-1474.

68

Methods of making *hedgehog* knock-out or disruption transgenic animals are also generally known. See, for example, *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Recombinase dependent knockouts can also be generated, e.g. by homologous recombination to insert recombinase target sequences flanking portions of an endogenous *hh* gene, such that tissue specific and/or temporal control of inactivation of a *hedgehog* allele can be controlled as above.

Exemplification

The invention, now being generally described, will be more readily understood by reference to the following examples, which are included merely for purposes of illustration of certain aspects and embodiments of the present invention and are not intended to limit the invention.

Example 1

Cloning and Expression of Chick Sonic Hedgehog

(i) Experimental Procedures

Using degenerate PCR primers, vHH50 (SEQ ID No:18), vHH30 (SEQ ID No:19) and vHH31 (SEQ ID No:20) corresponding to a sequence conserved between *Drosophila hedgehog* (Dros-HH)(SEQ ID No:34) (Lee, J.J. et al. (1992) *Cell* 71: 33-50; Mohler, J. et al., (1992) *Development* 115: 957-971) and mouse Indian *hedgehog* (*Ihh*) (SEQ ID No:10), a 220 base pair (bp) fragment was amplified from chicken genomic DNA. From 15 isolates, two distinct sequences were cloned, pCHA (SEQ ID No:35) and pCHB (SEQ ID No:36), each highly homologous to mouse *Ihh* (Figure 1). A probe made from isolate pCHA did not detect expression in embryonic tissues. Isolate pCHB, however, detected a 4 kb message in RNA prepared from embryonic head, trunk, or limb bud RNA. This cloned PCR fragment was therefore used as a probe to screen an unamplified cDNA library prepared from Hamburger Hamilton stage 22 (Hamburger, W. et al., (1951) *J. Morph.* 88: 49-92) limb bud RNA as described below.

A single 1.6 kilobase (kb) cDNA clone, pHH-2, was selected for characterization and was used in all subsequent analyses. The gene encoding for this cDNA was named *Sonic Hedgehog* (after the Sega computer game cartoon character). Sequencing of the entire cDNA confirmed the presence of a single long open reading frame potentially encoding for a protein of 425 amino acids (aa). The clone extends 220 bp upstream of the predicted initiator methionine and approximately 70 bp beyond the stop codon. No consensus polyadenylation signal could be identified in the 3' untranslated region. A second potential initiator

69

methionine occurs at amino acid residue 4. The putative translation initiation signals surrounding both methionines are predicted to be equally efficient (Kozak, M., (1987) *Nuc. Acids Res.* 15: 8125-8132). When the pHH-2 *Sonic* cDNA is used to probe a northern blot of stage 24 embryonic chick RNA, a single mRNA species of approximately 4 kb is detected in both limb and trunk tissue. The message size was predicted by comparing it to the position of 18S and 28S ribosomal RNA. Hybridized mRNA was visualized after a two day exposure to a phosphoscreen. Because the *Sonic* cDNA clone pHH-2 is only 1.6 kb, it is likely to be missing approximately 2.4 kb of untranslated sequence.

PCR Cloning

All standard cloning techniques were performed according to Ausubel et. al. (1989), and all enzymes were obtained from Boehringer Mannheim Biochemicals. Degenerate oligonucleotides corresponding to amino acid residues 161 to 237 of the *Drosophila hedgehog* protein (SEQ ID No:34) (Lee, J.J. et. al., (1992) *Cell* 71: 33-50) were synthesized. These degenerate oligonucleotides, vHH50 (SEQ ID No:18), vHH30 (SEQ ID No:19), and vHH3I (SEQ ID No:20) also contained Eco RI, Cla I, and Xba I sites, respectively, on their 5' ends to facilitate subcloning. The nucleotide sequence of these oligos is given below:

vHH50: 5'-GGAATTCCCAG(CA)GITG(CT)AA(AG)GA(AG)(CA)(AG)I(GCT)IAA-3'
 vHH30: 5'-TCATCGATGGACCCA(GA)TC(GA)AAICCGC(TC)TC-3'
 vHH3I: 5'-GCTCTAGAGCTCIACIGCIA(GA)IC(GT)IGC-3'

where I represents inosine. Nested PCR was performed by first amplifying chicken genomic DNA using the vHH50 and vHH30 primer pair and then further amplifying that product using the vHH50 and vHH3I primer pair. In each case the reaction conditions were: initial denaturation at 93° C for 2.5 min., followed by 30 cycles of 94° C for 45 s, 50° C for 1 min., 72° C for 1, and a final incubation of 72° C for 5 min. The 220 bp PCR product was subcloned into pGEM7zf (Promega). Two unique clones, pCHA (SEQ ID No:35) and pCHB (SEQ ID No:36) were identified.

DNA Sequence Analysis

Nucleotide sequences were determined by the dideoxy chain termination method (Sanger, F. et al., (1977) *Proc. Natl. Acad. Sci. USA* 74: 5463-5467) using Sequenase v2.0 T7 DNA polymerase (US Biochemicals). 5' and 3' nested deletions of pHH-2 were generated by using the nucleases Exo III and S1 (Erase a Base, Promega) and individual subclones sequenced. DNA and amino acid sequences were analyzed using both GCG (Devereux, J. et al., (1984) *Nuc. Acids Res.* 12: 387-394) and DNASTAR software. Searches for related

sequences were done through the BLAST network service (Altschul, S.F. et al., (1990) *J. Mol. Biol.* 215: 403-410) provided by the National Center for Biotechnology Information.

Southern Blot Analysis

Five (5) µg of chick genomic DNA was digested with Eco RI and/or Bam HI, fractionated on a 1% agarose gel, and transferred to a nylon membrane (Genescreen, New England Nuclear). The filters were probed with ³²P-labeled *hha* or *hhb* at 42°C in hybridization buffer (0.5% BSA, 500 mM NaHPO₄, 7% SDS, 1 mM EDTA, pH 7.2; Church, G.M. et al., (1984) *Proc. Natl. Acad. Sci. USA* 81: 1991-1995). The blots were washed at 63°C once in 0.5% bovine serum albumin, 50 mM NaHPO₄ (pH 7.2), 5% SDS, 1 mM EDTA and twice in 40 mM NaHPO₄ (pH 7.2), 1% SDS, 1mM EDTA, and visualized on Kodak XAR-5 film.

Isolation Of Chicken Sonic cDNA Clones

A stage 22 limb bud cDNA library was constructed in λgt10 using Eco RI/NotI linkers. Unamplified phage plaques (10⁶) were transferred to nylon filters (Colony/Plaque screen, NEN) and screened with α³²P-labelled pooled inserts from PCR clones pCHA (SEQ ID No:35) and pCHB (SEQ ID No:36). Hybridization was performed at 42° C in 50% formamide 2X SSC, 10% dextran sulfate, 1% SDS and washing as described in the Southern Blot procedure. Eight positive plaques were identified, purified and their cDNA inserts excised with EcoRI and subcloned into pBluescript SK+ (Stratagene). All eight had approximately 1.7 kb inserts with identical restriction patterns. One, pHH-2, was chosen for sequencing and used in all further manipulations.

Preparation Of Digoxigenin-Labeled Riboprobes

Plasmid pHH-2 was linearized with Hind III and transcribed with T3 RNA polymerase (for antisense probes) or with Bam HI and transcribed with T7 RNA polymerase according to the manufacturers instructions for the preparation of non-radioactive digoxigenin transcripts. Following the transcription reaction, RNA was precipitated, and resuspended in RNase-free water.

Whole Mount In Situ Hybridization

Whole-mount *in situ* hybridization was performed using protocols modified from Parr, B.A. et al.. (1993) *Development* 119: 247-261; Sasaki, H. et al. (1993) *Development* 118: 47-59; Rosen, B. et al. (1993) *Trends Genet.* 9: 162-167. Embryos from incubated fertile White Leghorn eggs (Spafas) were removed from the egg and extra-embryonic membranes dissected in calcium/magnesium-free phosphate-buffered saline (PBS) at room temperature. Unless otherwise noted, all washes are for five minutes at room temperature.

Embryos were fixed overnight at 4°C with 4% paraformaldehyde in PBS, washed twice with PBT (PBS with 0.1% Tween-20) at 4°C, and dehydrated through an ascending methanol series in PBT (25%, 50%, 75%, 2 X 100% methanol). Embryos were stored at -20°C until further use.

5 Both pre-limb bud and limb bud stage embryos were rehydrated through an descending methanol series followed by two washes in PBT. Limb bud stage embryos were bleached in 6% hydrogen peroxide in PBT, washed three times with PBT, permeabilized with proteinase K (Boehringer, 2 µg/ml) for 15 minutes, washed with 2 mg/ml glycine in PBT for 10 minutes, and twice with PBT. Pre-limb bud stage embryos were permeabilized (without prior incubation with hydrogen peroxide) by three 30 minute washes in RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% Deoxycholate, 0.1% SDS, 1mM EDTA, 50 mM Tris-HCl, pH 8.0). In all subsequent steps, pre-limb bud and limb bud stage embryos were treated equivalently. Embryos were fixed with 4% paraformaldehyde/0.2% glutaraldehyde in PBT, washed four times with PBT, once with pre-hybridization buffer (50% formamide, 5 X SSC, 15 1% SDS, 50 µg/ml total yeast RNA, 50 µg/ml heparin, pH 4.5), and incubated with fresh pre-hybridization buffer for one hour at 70°C. The pre-hybridization buffer was then replaced with hybridization buffer (pre-hybridization buffer with digoxigenin labeled riboprobe at 1 µg/ml) and incubated overnight at 70°C.

Following hybridization, embryos were washed 3 X 30 minutes at 70°C with solution 20 1 (50% formamide, 5 X SSC, 1% SDS, pH 4.5), 3 X 30 minutes at 70°C with solution 3 (50% formamide, 2 X SSC, pH 4.5), and three times at room temperature with TBS (Tris-buffered saline with 2 mM levamisole) containing 0.1% Tween-20. Non-specific binding of antibody was prevented by preblocking embryos in TBS/0.1% Tween-20 containing 10% heat-inactivated sheep serum for 2.5 hours at room temperature and by pre-incubating anti-25 digoxigenin Fab alkaline-phosphatase conjugate (Boehringer) in TBS/0.1% Tween-20 containing heat inactivated 1% sheep serum and approximately 0.3% heat inactivated chick embryo powder. After an overnight incubation at 4°C with the pre-adsorbed antibody in TBS/0.1% Tween-20 containing 1% sheep serum, embryos were washed 3 X 5 minutes at room temperature with TBS/0.1% Tween-20, 5 X 1.5 hour room temperature washes with 30 TBS/1% Tween-20, and overnight with TBS/1% Tween-20 at 4°C. The buffer was exchanged by washing 3 X 10 minutes with NTMT (100mM NaCl, 100 mM Tris-HCl, 50 mM MgCl₂, 0.1% Tween-20, 2 mM levamisole). The antibody detection reaction was performed by incubating embryos with detection solution (NTMT with 0.25 mg/ml NBT and 0.13 mg/ml X-Phos). In general, pre-limb bud stage embryos were incubated for 5-15 hours and limb bud stage embryos 1-5 hours. After the detection reaction was deemed complete, 35 embryos were washed twice with NTMT, once with PBT (pH 5.5), postfixed with 4% paraformaldehyde/0.1% glutaraldehyde in PBT, and washed several times with PBT. In

some cases embryos were cleared through a series of 30%, 50%, 70%, and 80% glycerol in PBT. Whole embryos were photographed under transmitted light using a Nikon zoom stereo microscope with Kodak Ektar 100 ASA film. Selected embryos were processed for frozen sections by dehydration in 30% sucrose in PBS followed by embedding in gelatin and freezing. 25 μ m cryostat sections were collected on superfrost plus slides (Fisher), rehydrated in PBS, and mounted with gelvatol. Sections were photographed with Nomarski optics using a Zeiss Axiophot microscope and Kodak Ektar 25 ASA film.

(ii) *Sequence Homology Comparison Between Chicken Sonic hh And Dros-HH And Other Vertebrate Sonic hh Proteins*

The deduced *Sonic* amino acid sequence (SEQ ID No:8) is shown and compared to the *Drosophila hedgehog* protein (SEQ ID No:34) in Figure 2. Over the entire open reading frame the two proteins are 48% homologous at the amino acids level. The predicted *Drosophila* protein extends 62 aa beyond that of *Sonic* at its amino terminus. This N-terminal extension precedes the putative signal peptide (residues 1-26) of the fly protein (SEQ ID No:34), and has been postulated to be removed during processing of the secreted form of *Drosophila hedgehog* (Lee, J.J. et al., (1992) *Cell* 71: 33-50). The sequence of residues 1-26 of the *Sonic* protein (SEQ ID No:8) matches well with consensus sequences for eukaryotic signal peptides (Landry, S.J. et al., (1993) *Trends. Biochem. Sci.* 16: 159-163) and is therefore likely to serve that function for *Sonic*. Furthermore, Figure 3 shows a hydropathy plot (Kyte, J. et al., (1982) *J. Mol. Biol.* 157: 133-148) indicating that residues 1-26 of the *Sonic* protein (SEQ ID No:8) exhibit a high hydrophobic moment in accord with identified eukaryotic signal peptides. Cleavage of the putative signal sequence should occur C-terminal to residue 26 according to the predictive method of von Henjje, G. (1986) *Nucl. Acid. Res.* 11: 1986. A single potential N-linked glycosylation site is located at amino acid residue 282 of the *Sonic* protein (SEQ ID No:8). The predicted *Sonic* protein does not contain any other strong consensus motifs, and is not homologous to any other proteins outside of the *Hedgehog* family.

The mouse (SEQ ID No:11) and zebrafish (SEQ ID No:12) homologs of *Sonic* have also been isolated. A comparison of these and the *Drosophila* sequence is shown schematically in Figure 4. All of the vertebrate proteins have a similar predicted structure: a putative signal peptide at their amino terminus, followed by an extraordinarily similar 182 amino acid region (99% identity in chicken versus mouse and 95% identity in chicken versus zebrafish) and a less well conserved carboxy-terminal region.

(iii) *At Least Three Hedgehog Homologues Are Present In The Chicken Genome*

Since two distinct PCR products encoding for chicken *hedgehogs* were amplified from genomic DNA, the total number of genes in the chicken *hedgehog* family needed to be

estimated. The two PCR clones pCHA (SEQ ID No:35) and pCHB (SEQ ID No:36) were used to probe a genomic Southern blot under moderately stringent conditions as described in the above Experimental Procedures. The blot was generated by digesting 5 µg of chick chromosomal DNA with EcoRI and BamHI alone and together. Each probe reacted most strongly with a distinct restriction fragment. For example, the blot probed with pCHA, shows three bands in each of the Bam HI lanes, one strong at 6.6 kb and two weak at 3.4 and 2.7 kb. The blot probed with pCHB, shows the 2.7 kb band as the most intense, while the 3.4 and 6.6 kb bands are weaker. A similar variation of intensities can also be seen in the Bam HI/Eco RI and EcoRI lanes. Exposure times were 72 hr. This data indicates that each probe recognizes a distinct chicken *hedgehog* gene, and that a third as yet uncharacterized chicken *hedgehog* homolog exists in the chicken genome.

(iv) *Northern Analysis Defining Sites Of Sonic Transcription*

Northern analysis was performed which confirmed that *Sonic* is expressed during chick development. The spatial and temporal expression of *Sonic* in the chick embryo from gastrulation to early organogenesis was determined by whole mount *in situ* hybridization using a riboprobe corresponding to the full-length *Sonic* cDNA (SEQ ID No:1).

20µg total RNA isolated from stage 24 chick leg buds or bodies (without heads or limbs) was fractionated on a 0.8% agarose formaldehyde gel and transferred to a nylon membrane (Hybond N, Amersham). The blot was probed with the 1.6 kb EcoRI insert from pHH-2. Random-primed α³²P-labelled insert was hybridized at 42°C hybridization buffer (1% BSA, 500mM NaHPO₄, 7% SDS, 1 mM EDTA, pH 7.2) and washed at 63° C once in 0.5% bovine serum albumin, 50 mM NaHPO₄ (pH 7.2), 5% SDS, 1 mM EDTA and once in 40 mM NaHPO₄ (pH 7.2), 1% SDS, 1mM EDTA. The image was visualized using a phosphorimager (Molecular Dynamics) and photographed directly from the video monitor.

(v) *Expression Of Sonic During Mid-Gastrulation*

Sonic message is detected in the gastrulating blastoderm at early stage 4, the earliest stage analyzed. Staining is localized to the anterior end of the primitive streak in a region corresponding to Hensen's node. As gastrulation proceeds, the primitive streak elongates to its maximal cranial-caudal extent, after which Hensen's node regresses caudally and the primitive streak shortens. At an early point of node regression, *Sonic* mRNA can be detected at the node and in midline cells anterior to the node. By late stage 5, when the node has migrated approximately one-third of the length of the fully elongated primitive streak, prominent *Sonic* expression is seen at the node and in the midline of the embryo, reaching its anterior limit at the developing head process. Sections at a cranial level show that *Sonic* mRNA is confined to invaginated axial mesendoderm, tissue which contributes to foregut and notochord. More caudally, but still anterior to Hensen's node, staining of axial

mesoderm is absent and *Sonic* expression is confined to the epiblast. At the node itself, high levels of *Sonic* message are observed in an asymmetric distribution extending to the left of and posterior to the primitive pit. This asymmetric distribution is consistently observed (6/6 embryos from stages 5-7) and is always located to the left of the primitive pit. At the node, and just posterior to the node, *Sonic* expression is restricted to the epiblast and is not observed in either mesoderm or endoderm. The expression of *Sonic* in the dorsal epiblast layer without expression in underlying axial mesoderm contrasts markedly with later stages where *Sonic* expression in underlying mesoderm always precedes midline neural tube expression.

10 (vi) *Expression Of Sonic During Head Fold Stages*

During the formation and differentiation of the head process, *Sonic* mRNA is detected in midline cells of the neural tube, the foregut, and throughout most of the axial mesoderm. At stage 7, *Sonic* message is readily detected asymmetrically at the node and in ventral midline cells anterior to the node. The rostral limit of *Sonic* expression extends to the anterior-most portions of the embryo where it is expressed in the foregut and prechordal mesoderm (Adelmann, H.B., (1932) *Am. J. Anat.* 31, 55-101). At stage 8, expression of *Sonic* persists along the entire ventral midline anterior to Hensen's node, while the node region itself no longer expresses *Sonic*. Transverse sections at different axial levels reveal that at stage 8 *Sonic* is coexpressed in the notochord and the overlying ventromedial neuroectoderm from anterior to Hensen's node to the posterior foregut. The levels of *Sonic* message are not uniform in the neural tube: highest levels are found at the presumptive mid- and hindbrain regions with progressively lower levels anterior and posterior. The increasing graded expression in the neural tube from Hensen's node to the rostral brain may reflect the developmental age of the neuroectoderm as differentiation proceeds from posterior to anterior. At the anterior-most end of the embryo, expression is observed in midline cells of the dorsal and ventral foregut as well as in prechordal mesoderm. Although the prechordal mesoderm is in intimate contact with the overlying ectoderm, the latter is devoid of *Sonic* expression.

(vii) *Expression Of Sonic During Early CNS Differentiation*

30 At stages 10 through 14, *Sonic* expression is detected in the notochord, ventral neural tube (including the floor plate), and gut precursors. By stage 10, there is a marked expansion of the cephalic neuroectoderm, giving rise to the fore- mid- and hind-brain. At stage 10, *Sonic* mRNA is abundantly expressed in the ventral midline of the hindbrain and posterior midbrain. This expression expands laterally in the anterior midbrain and posterior forebrain. Expression does not extend to the rostral forebrain at this or later stages. Sections reveal that *Sonic* is expressed in the notochord, the prechordal mesoderm, and the anterior midline of the

foregut. Expression in the neuroepithelium extends from the forebrain caudally. In the posterior-most regions of the embryo which express *Sonic*, staining is found only in the notochord and not in the overlying neurectoderm. This contrasts with earlier expression in which the posterior domains of *Sonic* expression contain cells are located in the dorsal epiblast, but not in underlying mesoderm or endoderm. Midgut precursors at the level of the anterior intestinal portal also show weak *Sonic* expression.

At stage 14, expression continues in all three germ layers. The epithelium of the closing midgut expresses *Sonic* along with portions of the pharyngeal endoderm and anterior foregut. Ectoderm lateral and posterior to the tail bud also exhibits weak expression. At this stage, *Sonic* is also expressed along entire length of the notochord which now extends rostrally only to the midbrain region and no longer contacts the neuroepithelium at the anterior end of the embryo. Expression in head mesenchyme anterior to the notochord is no longer observed. In the neural tube *Sonic* is found along the ventral midline of the fore- mid- and hindbrain and posteriorly in the spinal cord. In the forebrain, expression is expanded laterally relative to the hindbrain. At midgut levels, expression of *Sonic* in the neural tube appears to extend beyond the floor plate into more lateral regions. As observed at stage 10, *Sonic* at stage 14 is found in the notochord, but not in the ventral neural tube in posterior-most regions of the embryo. When neuroectodermal expression is first observed in the posterior embryo, it is located in midline cells which appear to be in contact with the notochord. At later stages, expression continues in areas which show expression at stage 14, namely the CNS, gut epithelium including the allantoic stalk, and axial mesoderm.

(viii) Sonic Is Expressed In Posterior Limb Bud Mesenchyme

The limb buds initially form as local thickenings of the lateral plate mesoderm. As distal outgrowth occurs during stage 17, *Sonic* expression becomes apparent in posterior regions of both the forelimb and the hindlimb. Sections through a stage 21 embryo at the level of the forelimbs reveal that expression of *Sonic* in limb buds is limited to mesenchymal tissue. A more detailed expression profile of *Sonic* during limb development is discussed below in Example 3. Briefly, as the limb bud grows out, expression of *Sonic* narrows along the anterior-posterior axis to become a thin stripe along the posterior margin closely apposed to the ectoderm. Expression is not found at more proximal regions of the bud. High levels of *Sonic* expression are maintained until around stage 25/26 when staining becomes weaker. Expression of *Sonic* is no longer observed in wing buds or leg buds after stage 28.

76
Example 2

Mouse Sonic Hedgehog Is Implicated in the Regulation of CNS and Limb Polarity

(i) *Experimental Procedures*

Isolation Of Hedgehog Phage Clones

5 The initial screen for mammalian *hh* genes was performed, as above, using a 700bp PCR fragment encompassing exons 1 and 2 of the Dros-HH gene. Approximately one million plaques of a 129/Sv Lambda Fix II genomic library (Stratagene) were hybridized with an α ³²P-dATP labeled probe at low stringency (55°C in 6xSSC, 0.5%SDS, 5 x Denhardt's; final wash at 60°C in 0.5 x SSC, 0.1% SDS for 20'). Five cross hybridizing
10 phage plaques corresponding to the *Dhh* gene were purified. Restriction enzyme analysis indicated that all clones were overlapping. Selected restriction enzyme digests were then performed to map and subclone one of these. Subclones in pGEM (Promega) or Bluescript (Stratagene) which cross-hybridized with the Dros-HH fragment where sequenced using an ABI automatic DNA sequencer.

15 Mouse *Ihh* and *Shh* were identified by low stringency hybridization (as described above) with a chick *Shh* cDNA clone to one million plaques of an 8.5 day λ gt10 mouse embryo cDNA library (Fahrner, K. et al., (1987) *EMBO J.* 6: 1265-1271). Phage plaques containing a 1.8kb *Ihh* and 0.64 and 2.8kb *Shh* inserts were identified. Inserts were excised and subcloned into Bluescript (Stratagene) for dideoxy chain termination sequencing using
20 modified T7 DNA polymerase (USB). The larger *Shh* clone contained a partially processed cDNA in which intron splicing at the exon 1/2 junction had not occurred.

To screen for additional *Ihh* and *Shh* cDNA clones, an 8.5 day λ ZAPII cDNA library was probed at high stringency (at 65°C in 6xSSC, 0.5% SDS, 5 x Denhardt's; final wash at 65 °C in 0.1xSSC, 0.1% SDS for 30') with the *Ihh* and *Shh* mouse cDNA clones. No additional
25 *Ihh* clones were identified. However several 2.6kb, apparently full length, *Shh* clones were isolated. The DNA sequence of the additional 5' coding region not present in the original 0.64 and 2.8kb *Shh* clones was obtained by analysis of one of the 2.6kb inserts.

Northern Blot Analysis

Expression of *Shh* was investigated by RNA blot analysis using 20 μ g of total RNA
30 from adult brain, spleen, kidney, liver, lung, 16.5dpc brain, liver and lung; 9.5dpc to 17.5dpc whole embryo; 9.5dpc forebrain, midbrain and 10.5dpc brain. RNA samples were electrophoretically separated on a 1.2% agarose gel, transferred and u.v. crosslinked to Genescreen (DuPont) and probed with 2X10⁶ cpm/ml of an α ³²P-dATP labeled mouse *Shh* probe (2.8kb insert from λ gt 10 screen). Hybridization was performed at 42°C in 50%
35 formamide 5x Denhardt's, 5xSSPE, 0.1%SDS, 6.5% dextran, 200 μ g/ml salmon sperm DNA.

Final wash was at 55°C in 0.1xSSC, 0.1%SDS. The blot was exposed for 6 days in the presence of an intensifying screen.

In Situ Hybridization, β -Galactosidase Staining And Histological Analysis

Embryos from 7.25 to 14.5dpc were analyzed for either *Shh* or HNF-3 β expression by whole mount *in situ* hybridization to digoxigenin labeled RNA probes as described in Wilkinson, (1992) *In situ Hybridization: A Practical Approach*. Oxford; Parr et al., (1993) *Development* 119:247-261. The mouse *Shh* probe was either a 2.8kb or 0.6kb RNA transcript generated by T7 (2.8kb) or T3 (0.6kb) transcription of XbaI and HindIII digests of Bluescript (Stratagene) subclones of the original *Shh* cDNA inserts. The HNF-3 β probe was generated by HindIII linearization of a HNF-3 β cDNA clone (Sasaki, H. et al., (1993) *Development* 118: 47-59) and T7 polymerase transcription of 1.6kb transcript. Embryos were photographed on an Olympus-SZH photomicroscope using Kodak Ektachrome EPY 64T color slide film.

Sections through wild type and WEXP2-C*Shh* transgenic embryos were prepared and hybridized with ³⁵S-UIP labeled RNA probes (Wilkinson, D.G. et al., (1987) *Development* 99: 493-500). Sections were photographed as described in McMahon, A.P. et al., (1992) *Cell* 69: 581-595.

β Staining of WEXP2-lacZ embryos with β was performed according to Whiting, J. et al., (1991) *Genes & Dev.* 5: 2048-2059. General histological analysis of wildtype and WEXP2-C*Shh* transgenic embryos was performed on paraffin sections of Bouin's fixed embryos counterstained with hematoxylin and eosin. Histological procedures were as described by Kaufman, M.H. (1992) *The Atlas of Mouse Development*, London: Academic Press. Sections were photographed on a Leitz Aristoplan compound microscope using Kodak EPY 64T color slide film.

DNA Constructs For Transgenics

Genomic *Wnt-1* fragments were obtained by screening a λ GEM12 (Promega) 129/Sv mouse genomic library with a 375 bp *MluI*-*Bgl*III fragment derived from the fourth exon of the murine *Wnt-1* gene. One of the clones (W1-15.1) was used in this study.

As an initial step towards the generation of the pWEXP2 expression vector, W1-15.1 was digested to completion with restriction enzymes *Aat*II and *Clal*, and a 2774 bp *Aat*II-*Clal* fragment isolated. This fragment was ligated into *Aat*II and *Clal* cut pGEM-7Zf vector (Promega), generating pW1-18. This plasmid was digested with *Hind*II and ligated to annealed oligonucleotides *lac*1 (SEQ ID No:21) and *lac*2 (SEQ ID No:22) generating pW1-18S* which has a modified polylinker downstream of the *Clal* restriction site. This construct (pW1-18S*) was digested with *Clal* and *Bgl*III and ligated with both the 2.5 kb 3' *Clal* - *Bgl*III